Forecasting the Price of Wheat and Other Commodities

by

Stephan Pfaffenzeller, MA

Thesis submitted to the University of Nottingham
for the Degree of Doctor of Philosophy, May 2002

Acknowledgments

I would like to acknowledge the generous support received from the Ministry of Agriculture Fisheries and Food through their research studentship.

Furthermore, I am indebted to my supervisors, Prof. Anthony Rayner and Prof. Paul Newbold for their help and guidance. I am also grateful to Dr. Thae Hwan Kim for his valuable assistance and to Betty Dow for providing me with commodity price data.

Table of Contents -Overview

Chapter 1: The Development of Primary Product Prices -a Survey of the Literature 1
1.1. Introduction 1
1.2. The Role of Primary Commodities in the Development Process 2
1.3. A Survey of the Debate on the Empirical Evidence for a secular decline in the Commodity Terms of Trade of LDCs. 8
1.4. Conclusion 29

Chapter 2: Data Series and Methodology 38
2.1. Description of the Data Series Used 38
2.2. Basic Econometric Methodology 46
2.3. Conclusion 57

Chapter 3: Estimation Results for the Trend Coefficient 83
3.1. Inference on stationarity 83
3.2. Estimation Results for ARIMA models 85
3.3. Structural Breaks 93
3.4. Comparison with Other Studies 112
3.5. Conclusion 119

Chapter 4: Further Attempts at Assessing the Significance of Trend Coefficient Estimates 163
4.1. Simulated Data Series for Selected Commodities 164
4.2. A Priori Inference on Unit Roots on the Basis of a Stationarity Test. 186
4.3. Modified Testing Methods for a Deterministic Trend in First Order Autoregressive Time Series 195
4.4. An Alternative Testing Procedure for the Significance of a Trend Coefficient 213
4.5. Conclusion 232

Chapter 5: Forecasts of Average Annual Commodity Prices Relative to MUV 259
5.1. Forecast alternatives for the commodity series in the sample 259
5.2. Selecting forecast models for different commodities 303
5.3. Relative Price Forecasts 328
5.4. Conclusion 354

Chapter 6: Trend-Cycle Decompositions for Individual Commodity Price Series 370
6.1. Trends and volatility in commodity price series 370
6.2. Fundamentals of the Beveridge Nelson Decomposition Method 372
6.3. Trend cycle decomposition results for integrated series 376
6.4. Trend and cycle components in stationary and trend stationary series 385
6.5. Conclusion: The relative importance trend and cycle components in relative commodity price series 391

Chapter 7: Overall Conclusions 403

Bibliography 410

Table of Contents

Chapter 1: The Development of Primary Product Prices -a Survey of the Literature 1
1.1. Introduction 1
1.2. The Role of Primary Commodities in the Development Process 2
1.2.1. The changing role of the agricultural sector 2
1.2.2. Non agricultural primary commodity prices 3
1.2.3. Relative commodity price developments and the role of primary sectors in economic development 5
1.3. A Survey of the Debate on the Empirical Evidence for a secular decline in the Commodity Terms of Trade of LDCs. 8
1.3.1. The developing discussion of the Prebisch-Singer Hypothesis. 8
1.3.2. Terms of trade developments during the 1970s and 1980s 10
1.3.3. Long run studies of terms of trade movements 12
1.3.3.1. Estimation methods used 13
1.3.3.2. The implications of price volatility and of inference regarding the order of integration of time series 17
1.3.3.3. The impact of structural breaks 22
1.3.3.4. Disaggregated data series 26
1.4. Conclusion 29
Appendix I.i. The Change of Foreign Exchange Earnings in Response to a Change in Export Prices 32
Appendix I.ii. Data Sources Used in the Literature 34
Appendix I.iii. Summary of Results in Various Studies on Trends in Commodity Terms of Trade 36

Chapter 2: Data Series and Methodology 38
2.1. Description of the Data Series Used 38
2.1.1. The original data sets 38
2.1.2 Description of the Data Series 43
2.2. Basic Econometric Methodology 46
2.2.1. Modelling univariate time series. 46
2.2.2. Testing for unit roots 50
2.2.3. Model estimation and selection 53
2.2.4. Further considerations for the formulation of forecast models 55
2.3. Conclusion 57
Appendix II.i. Average Annual Commodity Price Series Deflated by MUV. 58

Appendix II.ii. Graphical Illustrations of the Relative Commodity Price Series 70

Chapter 3: Estimation Results for the Trend Coefficient 83
3.1. Inference on stationarity 83
3.2. Estimation Results for ARIMA models 85
3.2.1. Evaluating trend estimates by commodity groups 88
3.3. Structural Breaks 93
3.3.1 The impact on trend estimates 97
3.3.2. Evaluating the impact of outliers on trend estimates by commodity groups 99
3.3.3. Overall evidence on secular trends 110
3.4. Comparison with Other Studies 112
3.4.1. A comparison with the results obtained by Grilli and Yang (1988) 113
3.4.2. A comparison with the results obtained by Cuddington (1992) 114
3.4.3. A comparison with the results of León and Soto (1997) 116
3.5. Conclusion 119
Appendix III.i. Unit Root Test Results 120
Appendix III.ii. Estimation Results for Relative Primary Product Price Series in Levels -minimum SBC specifications 129
Appendix III.iii. Estimation Results for Relative Primary Product Price Series in First Differences -minimum SBC specifications 135
Appendix III.iv. Estimation Results for Relative Primary Product Price Series in Levels -minimum SBC specifications including dummies 141
Appendix III.v. Estimation Results for Relative Primary Product Price Series in First Differences -minimum SBC specifications including dummies. 148
Appendix III.vi. Normality Tests for Residuals from ARIMA Estimates of Relative Primary Commodity Prices with and without Dummy Variables. 155
Appendix III.vii. Confidence Intervals for Coefficient Estimates 158

Chapter 4: Further Attempts at Assessing the Significance of Trend Coefficient Estimates 163
4.1. Simulated Data Series for Selected Commodities 164
4.1.1. Simulation Methodology 164
4.1.2. Fitting ARIMA Models to the Generated Data Series. 166
4.2. A Priori Inference on Unit Roots on the Basis of a Stationarity Test. 186

4.3. Modified Testing Methods for a Deterministic Trend in First Order Autoregressive Time Series 195
4.3.1. OLS with adjusted t-ratios 196
4.3.2. Maximum Likelihood Estimation Results 201
4.3.3. Generalised Least Squares Estimators for the Trend Coefficient 203
4.3.4. Adjusting for estimation bias in first order autoregression 206
4.3.5. An alternative approach to a priori testing for the order of integration 208
4.4. An Alternative Testing Procedure for the Significance of a Trend Coefficient 213
4.4.1. Vogelsang's test for a trend coefficient 213
4.4.2. Results of the Vogelsang Test 218
4.4.3. Simulation evidence on the Vogelsang Test Statistics 220
4.5. Conclusion 232
Appendix IV.i. Estimation Output for ARIMA Models with Higher Parameterisations 238
Appendix IV.ii. Further Details on the Leybourne-McCabe Stationarity Test 249
Appendix IV.iii. ADF Test Results Obtained Using Maximum Likelihood Estimation 255
Appendix IV.iv. Estimation Results for the Price Series for Hides. 257

Chapter 5: Forecasts of Average Annual Commodity Prices Relative to MUV 259
5.1. Forecast alternatives for the commodity series in the sample 259
5.1.1. The forecast alternatives 260
5.1.2. Pre-testing and Forecast Performance 261
5.1.3. Differences in the forecast performance of alternative models 268
5.1.4. Simulation results for the forecast performance of various model alternatives 274
5.1.5. Diebold and Killian Revisited 299
5.2. Selecting forecast models for different commodities 303
5.2.1. Selected Forecast Models for Individual Price Series 305
5.3. Relative Price Forecasts 328
5.3.1. Price forecasts obtained 328
5.3.2. Comparison with Worldbank forecasts 348
5.3.3. Assessing in-sample forecasts 351
5.4. Conclusion 354
Appendix V.i. Ten Period Forecasts and Confidence Intervals 356
Appendix V.ii. Forecasts for Wheat Prices from the Difference Stationary Model Selected by SBC 361

Appendix V.iii. Estimation Results for Relative Primary Product Price Series in First Differences -Minimum SBC Specifications 365

Chapter 6: Trend-Cycle Decompositions for Individual Commodity Price Series 370
6.1. Trends and volatility in commodity price series 370
6.2. Fundamentals of the Beveridge Nelson Decomposition Method 372
6.3. Trend cycle decomposition results for integrated series 376
6.4. Trend and cycle components in stationary and trend stationary series 385
6.5. Conclusion: The relative importance trend and cycle components in relative commodity price series 391
Appendix VI.i. Further Graphs of Commodity Price Series and Trend or Permanent Components. 393
Appendix VI.ii. Persistence Results for Difference Stationary Models 397

Chapter 7: Overall Conclusions 403

Bibliography 410

Abstract

The long term behaviour of primary product prices has been a central issue underlying projections of commodity price series. Against the background of the Prebisch Singer Hypothesis, the presence, magnitude and direction of a secular trend in commodity price series have themselves become the subject of a long standing debate.

This study uses the individual commodity price series underlying the Grilli and Yang data set and, where possible, extends these data series up to 1998. Deflating primary commodity prices by the MUV index, the question of trend components in the time series is studied considering evidence from univariate models and allowing for trend stationary or integrated data series with drift. In this context the impact of serial correlation in finite samples and the impact of wrongly modelling a data series as integrated are considered in detail. Further evidence from a trend test developed by Vogelsang (1998) is also taken into account.

In selecting forecast models, the usefulness of unit root pre-testing is assessed allowing for interdependence between the inferred order of integration and the significance of the trend or drift coefficient estimate obtained. Projections from univariate models are obtained for a ten year horizon and Beveridge-Nelson trend cycle decompositions are computed to assess the importance of volatility surrounding the forecasts. It is found that with regards to the past behaviour of primary commodities as well as for the forecasts obtained, the trajectory of primary commodity prices relative to the price of developed country manufactures exports is not generally characterised by a downwards trend.

Chapter 1

The Development of Primary Product Prices

-a Survey of the Literature

Chapter 1: The Development of Primary Product Prices -a Survey of the Literature

1.1. Introduction

Any forecast of relative primary commodity prices which is based on an extrapolation of historical data series naturally depends on the way in which such a time series is modelled. While the short run dynamics are often accounted for in autoregressive moving average (ARMA) models of the residual, the characterisation of long term developments can crucially depend on the presence, direction and magnitude of a trend component in the data series. The question of the presence and sign of a deterministic trend in series of primary commodity prices relative to manufactured goods prices have been the object of a long standing controversy. During the time after the second World War this debate has been centred around the Prebisch Singer Hypothesis.

During the early 1950s, Raul Prebisch and Hans W. Singer questioned some of the then prevailing paradigms on the process of economic development (*cf.* Singer (1950) and Prebisch (1959)). The core arguments of Prebisch and Singer predicted a secular decline in the relative price of primary commodities and therefore in the net barter terms of trade of developing countries. The present chapter will survey the debate surrounding the evidence in favour of the presence of a deterministic trend against this background. The remainder of the chapter is organised as follows: in section 1.2 the Prebisch Singer Hypothesis (PSH) and its immediate theoretical background are briefly reviewed. Section

1.3 surveys the debate surrounding the empirical evidence to date and addresses some of the methodological developments. Section 1.4 concludes.

1.2. The Role of Primary Commodities in the Development Process

1.2.1. The changing role of the agricultural sector

It is a well documented fact that the agricultural sectors of modern developed economies have shown a general tendency to decline in relative importance, both in terms of their participation in national income as well as in terms of employment in these sectors (*cf.* Ingersent and Rayner (1999), Antle (1988) and Anderson (1987)). These developments are generally in line with what should be expected by what has become known as *Engel's Law*, *i.e.* the observation that income elasticities for food products -or for agricultural products more generally- take lower values at higher income levels and that as a consequence the income share of primary commodities should be expected to decline at higher *per capita* income levels (*cf.* Ingersent and Rayner *op. cit.*). *Engel's Law*[1] as such does not make predictions about expected relative price developments between the agricultural and the manufacturing sector, as indeed it is formulated for a *ceteris paribus* change under a constant price assumption. In so far as resources shift into the non-primary sector one should of course expect a fall in the relative price of primary commodities to provide the price incentive, at least under full employment conditions. Anderson (1987)[2] and Antle (1988) give a more detailed account of the driving forces underlying this

[1] In its most basic forms, of course, Engel's Law does little more than describe the falling share of individual households' income that is spend on agricultural food products. In this sense, links to the evolution of *per capita* income throughout the economy, and conclusions on relative price developments or demand for non-food agricultural products are extensions of the basic theory.

[2] Anderson (*op. cit.*) also shows how the relative price of agricultural products can increase in some extreme cases.

relative price change and show that technological change biased towards the industrial sector will tend to accentuate the relative price decline.

Interestingly, this decline in the relative price of primary commodities compared to manufactured goods is contrary to the price expectations originally postulated: Ricardo (1951) expected the relative price of agricultural commodities (he quotes corn as an example) to increase as successively less fertile farm land is brought into production. The relative price of manufactures on the other hand was expected to decrease as innovations improved productivity, thus putting downwards pressure on manufacturing prices. (This at least would be the outcome in a sufficiently competitive market.)

This original expectation on the development of relative prices of agricultural commodities is of course dependent on a number of restrictive assumptions. The underlying assumption on the prevailing supply conditions -with arable area acting as a binding constraint on agricultural production- have been compensated by large productivity increases (*e.g.* through the use of artificial fertilisers, tractors *etc.*[3]) in the agricultural sector. Demand conditions at higher levels of income such as the low income elasticities implied by *Engel's Law* and the low price elasticities of demand for agricultural products considered by Prebisch (1959) and Singer (1950) further provide the conditions that can justify expectations of falling relative primary commodity prices.

1.2.2. Non agricultural primary commodity prices

The supply conditions in non-agricultural primary commodity sectors -as in agriculture- are characterised by natural resource constraints as well as developments in production technology, affecting the importance of given

[3] On this point, see Ingersent and Rayner (*op.cit.* Ch.2) for examples in Europe and the USA.

resource constraints as production or extraction limits. What tends to be systematically different here is the nature of demand. Demand in non-agricultural sectors is often derived demand linked to industrial activities[4] and will therefore depend on activity levels in the relevant industries as well as on efficiency gains and raw material saving technological progress[5]. One crucial aspect of the long run development of demand for these products is therefore that the main price determinants are not as easily modelled in terms of income and population growth, but should be seen in the context of a larger number of factors which develop in a less predictable way.

Nevertheless, relatively low demand elasticities and a consistent improvement in the efficiency of raw material utilisation have been advanced as reasons for *a priori* expectations of a secular relative price decline of non-agricultural raw materials (*cf. e.g.* Borzenstein *et.al.* (1994)). However, this viewpoint may overlook other medium to long term developments which can dominate these trends or have done so in the past. Among these, the development of new industries or technologies as well as economic transition processes in individual countries can have a crucial impact on price developments. From a more backward looking perspective, events like the two world wars of the 20th century have had strong effects on world primary commodity markets over several years (see for example Worldbank (2000)).

---

[4] This study excludes oil prices throughout, since in the oil sector price developments are often influenced by OPEC supply decisions.

[5] A number of authors have commented on this point. For Copper markets see for example Vial (1992) and Worldbank (2000) for more a general account of commodity price developments. Bloch and Sapsford (2000) point out that a sustained downward trend in relative commodity prices can be obscured by temporary accelerations of industrial growth.

1.2.3. Relative commodity price developments and the role of primary sectors in economic development.

Even though the primary sectors of an economy are unlikely to provide sufficient potential for the development process in the long run, they can play a crucial role in mobilising resources in so far as they can secure foreign exchange earnings that facilitate the acquisition of modern production technology and machinery. Export financed domestic investment can be seen as desirable -and as more desirable than foreign investment- not least because it does not give rise to the issue of future profit repatriation. It may also be favoured for political reasons, for example in view of the ownership and control of domestic industrial facilities.

If primary sector based export earnings are to provide a basis for economic development, then the secular development of export prices relative to import prices are clearly an issue of concern. Where the trade relationship of interest is one between primary commodity exporting developing countries and industrialised economies exporting manufactured products, this reduces to a general concern over the development of primary commodity prices relative to manufactured goods prices.

The Prebisch-Singer Hypothesis: The Prebisch-Singer Hypothesis contradicts the classical prediction of an improvement in the relative price of primary commodities. Singer (1950) and Prebisch (1959) focus on a situation in which developing countries as a group are mainly dependent on primary commodity exports (*i.e.* their sector of comparative advantage) while developed countries as a group export mainly manufactured products. The critique which has become known as the Prebisch-Singer Hypothesis postulates that the gains from trade are distributed disproportionately in favour of developed countries as a consequence of a bias in the terms of trade that favours manufactured commodities. According to this hypothesis, this bias was present not only in the then prevailing terms of trade but also in their development over time[6]. Prebisch quoted British terms of trade data in support of this latter point (as mentioned in Spraos (1980)). The theoretical explanation for this phenomenon rests upon the assumption that primary commodities are traded in competitive international markets where they face low price and income elasticities (Singer (1950), Myrdal (1989) and Prebisch (1959)) while the market for manufactured goods faces higher demand elasticities and is seen as far less competitive and therefore dominated by mark-up pricing (*cf.* also: Bloch and Sapsford (1997)). This process is reinforced by raw material saving technological innovations in developed economies which further reduce the demand for developing countries' exports. Low price elasticities for primary commodity exports immediately imply that it is difficult to improve the income terms of trade by increasing the trade volume[7].

Singer (1950) observed that international trade and the incentives arising from the operation of market forces are unlikely to provide a basis for domestic investment in industrial development. It is assumed that such investment needs

---

[6] As pointed out by Gandolfo (1994) it is necessary for a trade equilibrium, in addition to the comparative cost advantage itself, that the terms of trade lie within $(P_{LDC}/M_{LDC})<(P_T/M_T)<(P_{DC}/M_{DC})$, where P and M are the prices of primary commodities and manufactures respectively and the subscripts indicate the price in developing countries (LDC) developed countries (DC) and in international trading relations (T) respectively. This condition implies in principle that the point to which the relative price can fall at any point in time is at least bounded by the domestic price ratio in the economy with a comparative advantage in the production of primary products, *i.e.* the developing country.

[7] Net Barter Terms of Trade are here understood to be a ratio of export to import prices (primary product prices deflated by manufactured goods prices in the general case underlying the PSH). Income Terms of Trade are here defined as the total value of exports deflated by an index of import prices. The income terms of trade clearly depend on export volumes as well as relative prices.

to be undertaken on the basis of foreign currency earnings[8], not least to enable the purchase of foreign made capital goods. The availability of foreign exchange is in turn assumed to be mainly a function of trade so that the change in a period's foreign exchange earnings following a fall in the price of exports can be expressed, in very simple terms, as[9]:

[1.3.1] $$dY_r = dP_x X(1 + \eta_x),$$

where dY_r is the change in foreign exchange earnings, dP_x is the change in the price of exports, X the quantity of exports and η_x the demand elasticity for exports. Recalling that η_x is negative, it is then obvious that a fall in P_x can be expected to lead to a fall in foreign exchange (and export) earnings if $|\eta_x|<1$. A low demand elasticity for primary commodities therefore tends to reinforce the symptoms of the secular decline in primary commodity prices as well as the impact of temporary price fluctuations.

In an economy which is dependent on trade in primary commodities a significant surplus of foreign exchange over subsistence expenditure and associated imports for consumption is therefore most likely to be available during a cyclical rise in the relative price of primary commodities. (It should be remembered that under the Prebisch Singer Hypothesis such relative price increases are bound to be cyclical since in the long run the prevailing market structure and raw material saving technical progress produce a secular decline in developing countries' terms of trade.) This characteristic of foreign exchange availability would imply that the means for investment are available precisely

---

[8] Singer (1950) is aware of the possible role for foreign direct investment, but argues that foreign investment projects fail to integrate with the domestic economy in developing countries, and therefore don't contribute to the local industrial infrastructure.

[9] See Appendix I.i for the derivation of this expression.

when price incentives temporarily favour investment in primary sectors in spite of the fact that those sectors are in long run decline (Singer (1950)).

The conclusions concerning the limited potential of the primary sector and the market's inability to provide foreign exchange for domestic industrial investment led to strong policy conclusions favouring development strategies based on import substitution. (Myrdal (1989) and Prebisch (1959) were particularly outspoken in their advocacy of these policy measures). If insights into the development of primary commodity prices are to form the basis of any decision on policy interventions, a well established conclusion on the available empirical evidence is needed. The following section will look into the ongoing debate on the evolution of developing countries' terms of trade and primary commodity prices.

1.3. A Survey of the Debate on the Empirical Evidence for a secular decline in the Commodity Terms of Trade of LDCs.

1.3.1. The developing discussion of the Prebisch-Singer Hypothesis.

The Prebisch-Singer Hypothesis has been subject to criticism on various grounds from an early date. As in the case of inflation, price measurements in trade fail to account for quality improvements, a factor that may have contributed to the overestimation of prices of manufactured products (see for example Spraos (1980), Grilli and Yang (1988) or Bleaney and Greenaway (1993)). Another issue that has been raised is the development of transport costs. It has been alleged that a decline of transport costs could be at least partially responsible for the fall in the price of primary commodities, although this development is unlikely to have been a major determinant of terms of trade developments (Spraos 1980)[10]. One should also bear in mind that relative price developments merely provide information on the development of the barter terms of trade and it can be argued that the extent to which trade is beneficial for developing countries mainly depends on the income terms of trade rather than relative product prices. There is at least some evidence that in spite of low demand elasticities, and in spite of the fact that supply expansions seem to have been a contributing factor in the observed price developments, a number of developing countries have succeeded in stabilising their income terms of trade in spite of adverse developments in the barter terms of trade (Borzenstein and Reinhart (1994), Borzenstein *et. al.* (1994)). More generally, various authors -not least Singer (1958) and Singer (1975)- have focused on other characteristics of developing countries' economies. After his initial hypothesis on product price developments, Singer gradually shifted his focus towards the role of technological innovation and the impact of rising debt burdens on the economic performance of underdeveloped countries (Singer (1975)).

Other critics concentrated on issues of statistical measurement. The problems arising from aggregating across different primary products have been highlighted early by Kindleberger (1958) who pointed out that important differences exist between different product groups as well as between different geographical regions. This issue is still present in the ongoing discussion of the empirical evidence and is treated in some detail by Grilli and Yang (1988), Cuddington (1992) and León and Soto (1997) among others. Yet other critics

---

[10]More precisely, Spraos (1980) points out that commodity prices -if recorded at an international commodity exchange located in a developed country- tend to quote import prices c.i.f. and export prices f.o.b. In this case it is possible in principle, that a terms of trade deterioration for developing countries could be recorded when in fact a fall in transport costs has lead to lower prices.

questioned the data used by Prebisch in his original study and the estimates he arrived at. Prebisch used British terms of trade data which he regarded as representative of developed country terms of trade *vis á vis* developing countries (Spraos (1980)). Spraos (*op. cit.*) further elaborates on criticism on the ground that the data used are not as representative as supposed in Prebisch' study, and also tend to exaggerate the magnitude of the inferred negative trend in primary commodity prices.

Substantial research has also been undertaken into the relationship between relative commodity prices and developing countries' terms of trade. This is one of the issues highlighted early on in the debate by Kindleberger (1958). Among other studies, Grilli and Yang (1988), Bleaney and Greenaway (1993), Powell (1991), Lutz and Singer (1994) and Lutz (1999b.) have assessed the relationship between relative commodity prices and developing country terms of trade. León and Soto (1995) concentrate on the terms of trade of Latin American countries. Interestingly, León and Soto (1995) do not find empirical support for a secular decline of country specific terms of trade, although they confirmed the existence of a negative trend for 17 out of the 24 commodity price series in the Grilli and Yang data set (*cf.* León and Soto (1997)). A similar contrast is found in Grilli and Yang (1988). A detailed treatment of this topic is beyond the scope of this study, although some of the statistical issues involved are relevant here as well.

1.3.2. Terms of trade developments during the 1970s and 1980s

In their work, Prebisch and Singer were confined to using data up to the early post-war years and the question naturally arises whether the hypothesised sustained decline in the barter terms of trade has been observed in subsequent years. Views on this issue differ.

Spraos (1980) argues that an extended data set which covers the 1970s makes the hypothesis of a secular decline more difficult to sustain. Considering a yet larger time horizon, Maizels (1992) takes a different view, and states that developing countries' terms of trade recovered temporarily during the 1970s, but then fell sharply during the 1980s when the price deterioration reached dramatic proportions. This view of a continued deterioration in developing countries' barter terms of trade during the 1980s is also supported by Borzenstein *et. al.* (1994). A number of economic factors have been identified as causes of the decline in primary product prices. Demand for primary commodities in developed countries declined during the recession in the early 1980s as well as in consequence of raw material saving technological innovations (Maizels 1992). This decline in demand was reinforced by supply side developments: Borzenstein and Reinhart (1994) as well as Maizels (1992) point out that developing countries increased their supply of primary commodities to compensate for the fall in the barter terms of trade and in order to meet rising foreign denominated debt servicing requirements. The world supply of primary commodities further increased after 1989 when the transition economies of the former Soviet Union increased the volume of their primary sector exports (Borzenstein and Reinhart (1994)).

There seems to be at least some evidence on the general pattern of the terms of trade development during the 1970s and 1980s. It is desirable, though, to look into the data over a longer time horizon if one wishes to obtain information on the presence of a long run trend. If policy conclusions are to be drawn from the observed developments a more detailed investigation is also required. The following subsections provide a discussion of the results of a number of long term studies, which have used various approaches to time series modelling.

1.3.3. Long run studies of terms of trade movements

A major problem in investigating the evolution of primary product prices and developing countries' terms of trade arises from the lack of available continuous long run data series. One of the longest as well as the most frequently used data set in the present discussion is the one compiled by Grilli and Yang (1988)[11]. The Grilli and Yang Commodity Price Index (GYCPI) covers the period 1900-1986 using price data for 24 primary commodities and has recently been extended to cover the period up to 1992 (*cf.* León and Soto (1997))[12]. The index uses annual average price series from World Bank sources. To proxy developing countries' terms of trade, the GYCPI is usually deflated by the United Nations Manufacturing Unit Value index (MUV). This index is available over a long time period, although for the periods 1914-1920 and 1939-1947 the MUV had to be completed by interpolation (Grilli and Yang (1988)).

While several studies have extended the Grilli and Yang data set, Bleaney and Greenaway (1993), who also extended the data series to 1991, truncate the series before 1925 to remove the influence of inferred volatile price

---

[11] A comprehensive listing of the data sources and series used by various authors is given in Appendix I.ii, Appendix I.iii gives further details.

[12] Since the relevant commodity price data are updated regularly, the index can be extended further, at least for the majority of the commodities covered.

movements[13] around 1920-1921 and higher than normal commodity prices during the 1920s.

Shorter intervals are targeted by Reinhart and Wickham (1994) who work with quarterly data for a total of four primary commodity groups, although the corresponding data series is available only for the period 1957-1993. Borzenstein and Reinhart (1994) also use quarterly data, covering the period 1971-1992.

1.3.3.1. Estimation methods used

The most intuitively appealing approach towards measuring the size and presence of a secular downward trend in the net barter terms of trade obviously consists in simply regressing a relative price index on a linear trend using an estimation technique such as OLS or Maximum Likelihood estimation. In this case the estimating equation would be of the form:

[1.3.2] $$p_t = a + \beta t + u_t$$

where p_t is the price index, a a constant, β the coefficient on the time trend t, and u_t the error term. The difference stationary equivalent would be:

[1.3.3] $$(1 - L)p_t = \beta + v_t$$

where β now represents the drift term, L denotes the lag operator and v_t is the residual. In either case, dummy variables may be included to account for structural breaks and the error term is either taken to be white noise or is adjusted for serial correlation or more generally modelled as an autoregressive moving average process.

---

[13] One should recall here that León and Soto (1997) attribute the apparent break in the GYCPI around 1920/21 to an aggregation problem and that this phenomenon is not consistently observed at lower levels of aggregation or for individual price series.

This is the approach adopted by Spraos (1980), Sapsford (1985) and also by Grilli and Yang (1988). Spraos (*op. cit.*) and Sapsford (1985) do not test their series for the presence of unit roots although, given the early date of their studies, this should come as no surprise. Grilli and Yang (1988) do apply the Dickey Fuller test and find no evidence for the presence of unit roots. While Spraos (1980) finds no significant negative trend, Sapsford (1985) and Grilli and Yang (1988) do, providing estimates of -1.29% *p.a.* and -0.59% *p.a.* respectively. Some doubts about these estimates do, however, remain. Aside from the issue of structural instabilities discussed below, one should bear in mind the temporary recovery of primary commodity prices during the 1970s when assessing the results of Spraos (1980) and Sapsford (1985). Spraos used a data set providing data up to 1970 and obviously could not take account of the pronounced decline during the 1980s[14].

A number of studies based on structural models have been conducted by Bloch and Sapsford (*cf.* Bloch and Sapsford (1991/1992), Bloch and Sapsford (1997) and Bloch and Sapsford (2000)). In all these studies the Grilli and Yang commodity price index deflated by the MUV index for the period after 1948 is used to represent the commodity terms of trade. Bloch and Sapsford (1991/1992) and Bloch and Sapsford (1997) explain the overall development of primary commodity prices in terms of Prebisch and Singer effects. In this context, the Prebisch effect is identified in terms of different price setting behaviour in developed and developing economies: Prices in developing countries or for developing country products are assumed to be set in competitive markets whereas prices in industrialised economies are set through

[14]Sapsford (1985) did extend some of the data used by Spraos up to 1980 and 1982.

mark up pricing. The Singer effect refers to differing pattern of productivity development with factor neutral technological change in the primary sector and raw material saving and labour saving technological change in the manufacturing sector. Bloch and Sapsford (1991/1992) and Bloch and Sapsford (1997) present empirical estimates in order to quantify the incidence of the underlying causal factors in the development of relative primary commodity prices. Bloch and Sapsford (2000) finally adopt a similar methodology and also emphasise the role of growth in the manufacturing sector. They conclude that the tendency towards an overall decline in primary commodity prices has been obscured by periods of faster than normal manufacturing growth after world war two. In all these studies the commodity and manufactured goods price series are identified as a difference stationary.

The majority of studies, however, is based on pure time series estimates of the barter terms of trade. Among these studies, Ardeni and Wright (1992) and Sapsford *et. al.* (1992) find evidence of a deterministic downward trend in the commodity terms of trade. Ardeni and Wright (*op.cit.*) put their estimate at between -0.14% to -1.6% *p.a.* While Sapsford *et. al.* (1992) estimate a downward trend of between -0.2% and -0.7% *p.a.* Helg (1991) includes dummies for a segmented trend and finds a small but significant negative trend after 1920. León and Soto (1997) arrive at trend estimates of around -1.5% for those commodities where negative trends have been found, although the point estimates for some price series differ substantially.

A cointegrated model is estimated by Lutz (1999a.) who estimates an overall decline of 0.89% *p.a.* for the overall commodity price index excluding fuel, a decline of 0.43% and 0.44% for the food and non food agricultural product indices respectively and an average annual decline of 0.88% for metals[15]. Powell (1991) also estimates a cointegrated model and accounts for structural breaks by introducing a variable that absorbs the cumulative impact of fluctuations in structural breaks. On this basis, Powell (1991) attributes the decline in relative commodity prices to three discrete shocks in 1921, 1938 and 1975. von Hagen (1989) also opts for a cointegrated representation as an alternative to a trend stationary model and fails to reject the null hypothesis of a zero trend coefficient.

Cuddington and Urzúa (1989) estimate a downwards trend of -0.6% *p.a.*, however, this can only be sustained if a structural break in 1920/21 is ignored, and the drift coefficient for the difference stationary alternative model is not significant. Newbold and Vougas (1996) are also less certain about the evidence in favour of a deterministic trend. They emphasise the uncertainty surrounding the issue of unit roots and the importance of conclusions about the presence of unit roots for subsequent inference about the significance of a trend coefficient estimate. Similar results are reported by Newbold *et.al.* (2000) for two commodity price series (Wheat and Maize)[16]. Bleaney and Greenaway (1993) finally attribute most of the decline in primary commodity prices to a one-off drop in 1980, if high commodity prices prior to 1925 are excluded from the sample. Grilli and Yang (1988), Helg (1991) and Sapsford *et. al.* (1992) test for and reject the presence of unit roots.

---

[15] It is certainly interesting that the decline of the overall index exceeds that of any of its component parts. This may lead one to doubt the extent to which such a composite index is representative of individual commodity price series.

[16] A bivariate model has also been considered by Newbold *et.al.* (2000). In this case conclusions on the presence of a trend vary depending on the deflator used.

1.3.3.2. The implications of price volatility and of inference regarding the order of integration of time series

Real commodity prices have been characterised by sustained volatility as well as by an overall decline in the price level. It therefore appears relevant to consider both of these characteristics. If one wishes to respond to this change in price pattern with targeted policy interventions, the appropriate measures in response to an increase in price volatility as opposed to a sustained price decline would, of course, be different.

Where a secular decline may have seemed to call for long term strategic industrial policy measures, the presence of short term fluctuations would be more likely to motivate short term measures such as the use of stabilisation funds to insulate earnings from primary commodity exports from exogenous price fluctuations. The adequacy of this kind of intervention does depend on a number of conditions. When price interventions are undertaken to stabilise private earnings, one needs to act on the assumption that private individuals are not sufficiently capable of making the required intertemporal expenditure adjustments unaided. As Borzenstein *et. al.* (1994) argue, the ability of private agents in this respect has frequently been underestimated, in particular with respect to developing countries, and the welfare improvements resulting from such a policy move may therefore appear dubious. Another motive for stabilising intervention can be the desire to stabilise foreign exchange earnings on a national level, or to stabilise the income of public entities which are directly dependent on these export earnings. In such a case too, the usefulness of stabilising policy interventions would still be dependent on the long run behaviour of primary commodity prices. A number of studies by the IMF

(Cashin *et.al.* (1999b), Cashin and Patillo (2000) and Cashin and McDermott (2001)) have investigated the trajectory of commodity prices focusing mainly on variations in the price series. Cashin and McDermott (2001) in particular conclude that volatility rather than a consistent decline is the main characteristic of the intertemporal behaviour of primary commodities. Other studies, discussed below, have addressed the issues of shock persistence and volatility in the context of studies of deterministic or stochastic trend components.

If the observed volatility merely represents fluctuations around a stable mean, stabilisation policies can, at least in principle, be suitable in compensating the effects of transitory price instability, so long as price shocks are sufficiently short lived. If however the fluctuations occur around a secular trend the situation is likely to be different. If stabilisation measures in such a case aim at fixing earnings or prices at a stable level, they are bound to be ineffective in the long run. Furthermore, where fluctuations are indeed merely a short term phenomenon, stabilisation measures which enforce mean reversion can be expected to have only a short term impact on the market process -even where, and in so far as, they distort prevailing price signals. If the presence of a secular trend is disregarded, however, a stabilising policy strategy would aim at fixing either prices or earnings at a level which is increasingly diverging from the equilibrium position. This would not only lead to a misallocation of public funds which is likely to be unsustainable in the long run but may also establish persistent distortions of investment incentives and support prolonged investment in sectors which are in long run decline.

In assessing the stationarity or trend characteristics of a time series one should therefore consider not only the presence of unit roots and take care to distinguish difference stationary processes from trend stationary processes. It is also essential to identify drift terms in difference stationary models, as these would indicate that in the case of mean reverting shocks the price level they would revert to does not remain constant. (In other words, mean reversion is defined with respect to the mean rate of change rather than some equilibrium price level.) Some authors go further than this and attempt to distinguish the permanent and automatically mean reverting components of stochastic shocks. Such a distinction would be relatively easy to make in a trend stationary time series where cyclical disturbances can be identified as temporary deviations from the deterministic trend. If the time series is subject to a stochastic trend, however, the issue is more complex.

A number of authors such as Cuddington and Urzúa (1989), Cuddington (1992) and Reinhart and Wickham (1994)[17] use the Beveridge-Nelson decomposition method to distinguish the cyclical and permanent components of a time series following a stochastic trend. Following this approach, the permanent component of a stochastic trend is the part of an innovation which can be assumed to persist in the absence of further shocks. The cyclical component in contrast would be expected to die out after a number of periods. In the studies mentioned, the authors employ trend-cycle decompositions in order to assess the scope for stabilising policy measures. Cuddington (1992:*217*) states explicitly that: "*A good understanding of the cyclical behavior of commodity prices is ... essential when considering countercyclical stabilization policies ...*"

[17] Reinhart and Wickham (1994) also use the Kalman filter.

Reinhart and Wickham also emphasise that instruments such as stabilisation funds are better suited to dealing with transitory disturbances of short duration. Accordingly, they focus on the cyclical behaviour and volatility characteristics of relative commodity prices in their study.

León and Soto (1997) criticise the use of the conventional persistence measures usually employed with the Beveridge and Nelson method and use a variance ratio statistic to assess persistence instead. León and Soto (1997) argue that the conventional parametric estimate of shock persistence tends to overstate the duration of shocks. Inferring shorter shock duration from the variance ratio statistic, they point to the understated scope for stabilising intervention, if the Beveridge Nelson decomposition is relied on. It is not generally clear though which mistake is more costly. The cost of misguided stabilising intervention is likely to accrue in form wasted exchange reserves and possibly accumulating public sector debts. The cost of precipitated adjustment to a transitory shock perceived as permanent would consist in the adjustment costs, which would differ by sector and in function of local market conditions. Moreover, the time period considered for mean reversion is quite large. (León and Soto (1997) recommend a period of up to half the sample size, to evaluate shock persistence.)

The presence of unit roots in time series on primary commodity prices is of importance not only in drawing conclusions on the scope for stabilising policy intervention, but also when estimating the presence and magnitude of a secular trend. It is also important to consider in this context that conclusions concerning the presence of unit roots are not always possible with a high degree of certainty. Newbold and Vougas (1996) and Newbold *et. al.* (2000) show how conclusions on the presence of unit roots depend on relatively minor changes in the significance levels employed in tests. Newbold and Vougas (1996) find a significant negative trend of around -0.8% *p.a.* when estimating a trend stationary model of the relative price of primary commodities (specified as GYCPI/MUV). When estimating a difference stationary model, the trend coefficient estimate has a similar magnitude but is statistically insignificant. Similar results are reported in Newbold *et.al.* (2000) where a downward trend of approximately 0.9-1% per year is estimated for Wheat and Maize prices. These estimates appear significant in trend stationary but not in difference stationary models.

León and Soto (1997), contrasting their results with Cuddington (1992), identify a larger number of trend stationary series and also identify a larger number of statistically significant trend coefficient estimates -most of them in trend stationary models.

Other authors are also somewhat cautious in their conclusions on the presence of unit roots. Cuddington and Urzúa (1989) estimate trend stationary models and difference stationary models. After applying Perron's test for a unit root in time series with changing mean, Cuddington and Urzúa conclude that a difference stationary model is more appropriate but nevertheless also estimate the trend stationary model to allow for comparisons with other studies that have proceeded similarly. In this case, the authors find no evidence of a secular downwards trend for either the difference stationary or the trend stationary specification so long as structural breaks are accounted for in the trend stationary model. (*i.e.* the trend coefficient estimate in the trend stationary model does appear significant if the dummy variable is not included. The drift coefficient estimate is shown to be insignificant whether or not a dummy variable is included in the difference stationary model.)

Sapsford *et.al.* (1992) follow the same unit root pre-testing procedure as Cuddington and Urzua (1989) for a modified data set. Based on their test result they opt for a trend stationary model and find evidence in favour of a significant trend coefficient estimate when a dummy variable is included in the trend stationary model.

What the studies surveyed above illustrate is that different conclusions on the presence of trend coefficients often seem to be strongly influenced by the assumed order of integration. A large number of studies base their analysis on *a priori* conclusions from unit root tests. The tests employed range from simple Dickey-Fuller tests (*e.g.* in the case of Grilli and Yang (1988)) to tests accommodating a structural break (*e.g.* in Cuddington (1992) and León and Soto (1997)). Ahrens and Sharma (1997) even approach the subject with a set of 4 different unit root tests in their study of US commodity prices[18]. While there does not seem to be a consensus as to which unit root test ought to be employed in studies of relative commodity prices, the importance of *a priori* conclusions regarding the order of integration is widely accepted.

1.3.3.3. The impact of structural breaks

Conclusions on the presence of a deterministic negative trend in real primary commodity prices can also depend on the impact of structural breaks. A related issue is the appropriate starting point for the data period under investigation, since mistaken conclusions on long run trends can be drawn if estimation is

[18] It appears though that Ahrens and Sharma take the presence of trends as given and merely attempt to distinguish stochastic and deterministic trends.

based on a data set with the earliest data points at unusually high or low levels. This issue is emphasised by Bleaney and Greenaway (1993) who truncate the Grilli and Yang data set before 1925 to account for high relative commodity prices before this date. The disagreement between Spraos (1980) and Sapsford (1985) is mainly based on the impact of a possible structural break in 1950. While Spraos (1980) rejects the hypothesis of the presence of a deterministic negative trend in the relative price of primary commodities, Sapsford (1985) shows that a significant downwards trend of -1.29 % *p.a.* can be identified if a structural break in 1950 is accounted for. Among the studies using the Grilli and Yang data set, the study by Cuddington and Urzúa (1989) highlights the importance of structural breaks. (Cuddington and Urzúa (*op.cit.*) find a significant downward trend only when disregarding the structural break in the data series in 1920/1921). The presence of this structural break also leads the authors to infer the presence of unit roots on the basis of Perron's test. This is crucial, since their estimates for a difference stationary model provide no evidence for a significant trend. In their critique of Cuddington and Urzúa (*op.cit.*), Sapsford *et. al.* (1992) focus on the extent of the structural break in 1920/1921. Using a modified data set with reduced structural instability at this point in time[19] they reject the difference stationary model and estimate a trend stationary model with a significant downwards trend. In yet another study based on the Grilli and Yang index Helg (1991) concludes that the presence of a trend can be confirmed but that the slope of the trend line changes after 1920.

[19] The MUV index used by Grilli and Yang (1988) had to be completed by interpolation for the periods 1914-1920 and 1939-1947, where data were missing. Sapsford *et.al.* (1992) use British terms of trade data to fill the 1914-1920 gap.

It has been mentioned above that Bleaney and Greenaway (1993) find support for a deterministic downward trend to be substantially weakened if data prior to 1925 are omitted from the sample. They further show that part of the remaining fall in relative commodity prices can be attributed to a structural break in 1980. Evidence in favour of a deterministic downward trend is further weakened if the period prior to 1980 is accounted for by a dummy variable.

Aside from this direct impact, the consequences of structural instabilities of time series data can become manifest in a more indirect fashion through the link between structural breaks and the validity of unit root test results. Conclusions on the presence of unit roots depend not only on the significance levels employed in unit root tests, but also on the presence of structural breaks in the data series in question. It has been established by Perron (1989) that conventional unit root tests have very low power in data series with a structural break. This implies that in those studies in which no structural breaks are found and conventional unit root test such as the augmented Dickey-Fuller (ADF) test are employed to test for unit roots, any conclusions on the presence of unit roots are conditional on the validity of the underlying assumptions on the presence of structural breaks. The issue of structural breaks is also of concern in the case of Ardeni and Wright (1992) who are aware of a possible structural break in 1921 and reject the unit root null hypothesis using the ordinary Dickey-Fuller test while noting that the Augmented Dickey-Fuller test fails to reject the null hypothesis of a unit root. They conclude that the evidence for non-stationarity is weak and show that their results are robust to estimating a trend stationary model for a sample truncated before 1922. A number of studies (Cuddington (1992), Cuddington and Urzúa (1989), Reinhart and Wickham

(1994) and Sapsford *et. al.* (1992)) use a modified unit root test developed by Perron (1989) to correct for the impact of structural breaks.

The issue of the presence of structural breaks in commodity price series is problematic, since there still seems to be substantial uncertainty concerning the presence as well as the number and position of structural breaks in the terms of trade data used. The presence of a structural break is often reported in 1920/1921, for the Grilli and Yang data set. This is not accepted by all authors, however. Ardeni and Wright (1992), León and Soto (1997) and Sapsford *et. al.* (1992) have doubts about the empirical relevance of this assumed structural break, though Sapsford *et. al.* (1992) are only able to mitigate the extent of the structural break in the original Grilli and Yang data set after substituting data for Britain's terms of trade for the period 1914-1920. Ardeni and Wright (1992) have doubts about the 1920/1921 structural break and finally resort to estimating a truncated data series (it will be recalled that a similar truncation was imposed by Bleaney and Greenaway (1993)). León and Soto (1997) aim to account for endogenously inferred structural breaks in their unit root testing procedure for individual commodity price series. Grilli and Yang find no significant evidence of a structural break while Newbold and Vougas (1996) confirm a sharp drop in the data series in 1920 but allow for the possibility of this representing a single outlier rather than a level shift in the data series. Sapsford (1985) and Sapsford *et. al.* (1992), finally, report a structural break in 1950.[20]

[20] Sapsford (1985) reports a structural break for 1950 rather than 1920/21 using a different data set from Sapsford *et.al.* (1992). Sapsford *et. al.* (1992) report the presence of structural breaks in 1950 as well as in 1920/21.

1.3.3.4. Disaggregated data series

While many of the studies reviewed above have treated aggregate commodity price and terms of trade series, there are a number of authors who considered a separate treatment of disaggregated data series. The inadequacy of aggregate terms of trade data series for an assessment of the development implications of international commodity price developments had been recognised as early as 1958 by Kindleberger (*op. cit.*)[21]. Among more recent treatments of the problem, Grilli and Yang (1988) considered the developments in the relative prices of a number of commodity groups. They distinguish four main commodity groups: 1. Food [GYCPIF] (and within this group tropical beverages and other food items), 2. non food agricultural commodities [GYCPINF] and 3. metals [GYCPIM]. They find an overall negative trend of -0.36% *p.a.* for food prices but within this group, prices for tropical beverages (*i.e.* Coffee, Tea and Cocoa) show a positive trend of around 0.63% *p.a.* while for other food[22] prices the trend is of -0.54% *p.a.* This is significant for LDCs in so far as tropical beverages are exported almost exclusively by LDCs while among the other food items developed countries' exports, and in particular developed countries' exports to LDCs, play a significant role (Grilli and Yang (*op. cit.*)). For non food agricultural commodities they find a negative trend of -0.84% *p.a.*, while the measured downwards trend for metal prices is of -0.82% *p.a.*[23]

---

[21] The assumption homogeneous time series behaviour of relative commodity price series is often referred to as comovement or excess comovement. This assumption of commodity price comovement was criticized more recently and in a different methodological context by Cashin *et. al.* (1999a.).

[22] They report a somewhat larger negative trend of -0.68% *p.a.* for cereals.

[23] In interpreting this result the authors observe that metal prices followed a positive trend for

Reinhart and Wickham (1994) infer negative trends for all three commodity groups (beverages, food and metals). Cuddington (1992) looks into the trends of the 26 individual primary commodity price series (including fuel prices) underlying the Grilli and Yang data set. He finds negative trends in four cases, five series are found to follow positive trends while the remaining 17 price series do not provide significant evidence of a trend in either direction. These results are in marked contrast with those obtained by León and Soto (1997). Using an extended version of the Grilli and Yang data set for 24 commodities, they infer the presence of a negative trend in 17 cases while four series seem to follow a positive trend. They find no evidence of a trend for three of the commodity price series.

As was the case for studies of aggregate price indices, the presence of unit roots is an issue in estimating individual commodity price series. Cuddington (1992) finds unit roots for 13 of the 26 price series considered. He accordingly estimates 13 difference stationary and 13 trend stationary models, with the results reported above. León and Soto, in contrast, find a smaller number of unit roots, and estimate 20 trend stationary and four difference stationary processes. Newbold *et. al.* (2000) follow the methodology in Newbold and Vougas (1996) when estimating the relative price trends of individual commodity price series for Wheat and Maize. Like Newbold and Vougas they find similar (negative) trend estimates for both commodities using either trend stationary or difference stationary models, although as in the case of the aggregate series used in Newbold and Vougas (*op. cit.*), there are large

---

the period 1942-1986 in spite of productivity improvements.

differences in the significance levels for the trend coefficients obtained for different models.

As for aggregate price series, structural breaks are again an issue for individual relative price series. In contrast to the case of aggregate primary commodity price series where a number of authors found structural breaks in either 1950 or 1920/21, results in the case of individual commodity price data differ much more. Grilli and Yang (1988) find a structural break in 1940 for their metal price index and report none for the other price indices. Cuddington (1992) finds structural breaks for only two individual price series: for coffee in 1950 and for oil in 1974. León and Soto (1997) report structural breaks in a number of different years, but none for 1921.

The structural break in 1920/1921 observed for the GYCPI has been frequently reported for the aggregate index but not for the sub-indices or individual relative price series. These results obviously cast some doubt on the usefulness of aggregate price indices in general. What implications a global trend in relative commodity prices would have for the development prospects of less developed countries depends crucially on whether certain commodities can generally be identified as exports from LDCs to DCs or the other way around. (This may be the case for tropical beverages and cereals respectively, as pointed out by Grilli and Yang *op. cit.*) In the case of individual developing economies one should however look into the development of particular price series for which the relevance for an individual country's trade balance can be assessed.

1.4. Conclusion

Various studies have been dedicated to the development of relative primary commodity prices. Among these studies, univariate models incorporating a linear trend are frequently used, as is the Grilli and Yang Commodity Price Index (GYCPI). In so far as there is an area of consensus in the use of univariate models, more recent contributions use trend stationary and difference stationary models allowing for autoregressive and, in some cases, moving average components in the residual process. In addition to the question of trend detection itself, there remains an interlinked set of controversial points comprising a.) the presence, number and position of structural breaks, b.) the suitability of composite indices or subindices and c.) the reliability of unit root pre-tests in this particular context. It is obvious from the survey of a number of studies which either use composite indices at different levels of aggregation (*e.g.* Grilli and Yang (1988) and León and Soto (1997)) as well as studies using individual commodity price series (*e.g.* Cuddington (1992), León and Soto (1997) and Newbold *et. al.* (2000)) that the recorded evidence on structural breaks changes with the level of aggregation. The structural break in 1920/21 which has been recorded frequently for the overall GYCPI, does not appear to be frequently observed when data series are disaggregated. It has further been observed that assumptions regarding the presence of structural breaks are linked to the reliability of unit root pre-tests.

Studies like Newbold *et.al.* (2000) and Newbold and Vougas (1996) further suggest that there may be a systematic link between inference regarding the order of integration of a data series and subsequent conclusions on the presence of a statistically significant trend or drift term. This impression is further strengthened by a comparison of the studies of Cuddington (1992) and León and Soto (1997). In both cases models are estimated for individual commodity price series. It is also the case though, that León and Soto (1997) identify a larger number of trend stationary models than Cuddington (1992), while simultaneously finding a larger number of significant trend estimates. Such a link would of course come as no surprise in the simple case of spurious rejections when a difference stationary series is modelled as trend stationary. A question that remains is whether standard errors are higher for drift coefficient estimates when modelling a trend stationary series as difference stationary.

Such an interdependence of conclusions regarding the trend or drift component and *a priori* assumptions on trend stationarity or difference stationarity would be of importance in its own right, since the mean reverting characteristics and trend movements of commodity price series can provide crucial background information for policy decisions. Moreover, conclusions regarding the presence of trend terms -or drift components in an integrated series- are clearly important for forecasts of the price series. This study will also attempt to identify the presence of trend or drift components in the series without relying on unit root test results as *a priori* assumptions. It will further be attempted to identify the most appropriate forecast model and obtain information on shock persistence, considering alternative options for the underlying order of integration of the data generating process.

The remainder of this study is structured as follows. Chapter 2 provides a general description of the data used, the transformations undertaken and of the general econometric methodology. Chapter 3 will investigate whether the presence of larger standard errors for the drift coefficient estimate -compared with the trend coefficient estimate in a trend stationary model- is a general characteristic of the constituent data series of the GYCPI. It will also be investigated whether the inclusion of possible structural breaks does substantially alter the results obtained. Chapter 4 provides further insights into the relationship between model specification and the conclusions reached on the presence of a trend term. It also shows the results of applying a testing procedure for trend coefficient estimates, developed by Vogelsang (1998), which is designed to be insensitive to misspecifications of the order of integration, so long as the data series is either I(0) or I(1). Forecast models are selected in Chapter 5 where the forecasts obtained are also presented. Chapter 6 presents trend cycle decomposition results for all the price series considered. Chapter 7 concludes.

Appendix I.i. The Change of Foreign Exchange Earnings in Response to a Change in Export Prices

Foreign currency revenues arising from trade can be written as follows:

I.i.i. $Y_r = rP_X X - P_M M,$

This is equal to the trade balance in foreign currency terms, where *Yr* denotes foreign currency earnings, *r* is the exchange rate, P_X the price and *X* the quantity of exports, P_M the price and *M* the quantity of imports. Without loss of generality, one can normalise $r = 1$. It is further assumed that *Px* and P_M are independent and that:

[I.i.ii.] $\frac{dM}{dP_X} = 0$

This assumption appears justified given the focus on countries specialising in the production and export of primary commodities. The rate of change of *Yr* in response to a change in *Px* is then:

[I.i.iii] . $$\frac{dY_r}{dP_X} = X + P_X \frac{dX}{dP_X}$$

Multiplying both sides of this equation by $\frac{X}{X} = 1$ gives:

[I.i.iv.] $$\begin{aligned} \frac{dY_r}{dP_X} &= X + \frac{dX}{dP_X}\frac{P_X}{X}X \\ &= X + X\eta_X \\ &= X(1 + \eta_X) \end{aligned}$$

Multiplying this by *dPx* gives the actual change in foreign currency earnings:

[I.i.v.] $dY_r = dP_X X(1 + \eta_X)$

(Another way of looking at the export price P_X is to interpret it as the relative price of exports with the import price set constant at P_M=1. Since the variable of interest is obviously some sort of real price, using an independent import price as deflator should be appropriate in this context. Formally one could substitute

[I.i.vi.] $$p_x = \frac{P_X}{P_M}$$

for P_X above, where obviously $p_x = P_X$ if $P_M = 1$. In the absence of a systematic relationship between export prices and import prices this relative price formulation should moreover improve the general applicability of the results presented here: an exogenous variation in the price of imports affects the purchasing power of export revenue. Under the relative price interpretation presented above, we have:

[I.i.vii.] $$p_x X = \frac{P_X X}{P_M}$$

i.e. the income terms of trade. (correspondingly, dY_r represents the change in the income terms of trade if p_x is substituted for P_X). The above results for a change in the export price would then continue to hold for a given level of import prices, while the (income and net barter) terms of trade would be affected by exogenous changes in P_M.

Appendix I.ii. Data Sources Used in the Literature

Table I.ii.i. Data sources used in various studies

Author (and year of publication)	**Data sources used**
Ahrens and Sharma (1997)	Commodity price data from various sources deflated by PPI
Ardeni and Wright (1992)	GYCPI and MUV
Bleaney and Greenaway (1993)	GYCPI deflated by MUV, data are extended to 1991 and truncated before 1925.
Bloch and Sapsford (2000)	GYCPI and MUV updated as in Bleaney and Greenaway (1993), for the period after 1948.
Bloch and Sapsford (1997)	GYCPI and MUV as in Grilli and Yang , for the period after 1948
Bloch and Sapsford (1991/92)	GYCPI and MUV as in Grilli and Yang, for the period after 1948
Borzenstein and Reinhart (1994)	Quarterly primary product price data from the IMF's International Financial Statistics (1971-1992), deflated by the US GNP deflator.
Cuddington (1992)	Data as for the GYCPI for 24 individual commodities and data for oil and coal prices, MUV as deflator
Cuddington and Urzúa (1989)	GYCPI and MUV as in Grilli and Yang
Grilli and Yang (1988)	Grilli and Yang Commodity Price Index (GYCPI) from annual US$ values of 24 primary commodities base weighted to their 1977-79 average. (Also: several variations on this, using different weights and one index using fuel prices). For manufactured goods: the UN index of Manufacturing Unit Values (MUV) completed for missing data periods. (Also: the US wholesale price index of industrial commodities USMPIO).
Helg (1991)	GYCPI deflated by MUV
León and Soto (1997)	GYCPI and MUV as in Grilli and Yang, updated to 1992 from IMF commodity price database.

Author (and year of publication)	**Data sources used**
Lutz (1999a.)	GYCPI deflated by MUV and updated to 1995 from IMF/IFS and UN Monthly Bulletin of Statistics
Newbold *et.al.* (2000)	Wheat and Maize price series as in GYCPI (WPI and MUV as deflator).
Newbold and Vougas (1996)	GYCPI and MUV as in Grilli and Yang
Powell (1991)	GYCPI deflated by MUV.
Reinhart and Wickham (1994)	Quarterly data from 1957-1993 for four primary commodity groups.
Sapsford (1985)	As Spraos (1980), extended coverage
Sapsford *et.al.* (1992)	GYCPI deflated by MUV, UK terms of trade data in place of some of the MUV.
Spraos (1980)	Variants of League of Nations data (annual 1921-1938), United Nations data (annual 1900-1970), World Bank and UNCTAD data (annual, from 1950 up to 1977).
von Hagen (1989)	GYCPI deflated by MUV

Note: in the case of studies based on structural models, only the data sources for commodity price series are given here.

Appendix I.iii. Summary of Results in Various Studies on Trends in Commodity Terms of Trade

Table I.iii.i. Overview of findings on trends in relative primary commodity price series.

Authors	Model/Regressors	Trend-Cycle Decomposition	Trend	Unit Roots	Index Decomposition	Structural Breaks
Ahrens and Sharma (1997)	ARMA	no	yes	yes some	Individual series no index	yes
Ardeni and Wright (1992)	ARMA	no	yes	no	no	yes
Bleaney and Greenaway (1993)	Regression of first difference on lagged level and trend.	no	no	possible	yes	yes
Cuddington (1992)	ARMA	Yes (BNDC)	yes, some	yes, some	yes, 26 individual price series	yes, in some series
Cuddington, Urzúa (1989)	ARMA	Yes (BNDC)	yes (without breaks)	yes	no	yes
Grilly and Yang (1988)	TS	no	yes	no	yes (commodity groups)	no
Helg (1991)	Regression on trend and dummies	no	yes	no	no	yes

Authors	Model/Regressors	Trend-Cycle Decomposition	Trend	Unit Roots	Index Decomposition	Structural Breaks
León and Soto (1997)	ARMA	Yes, variance ratio	yes	yes, some	yes (individual series)	yes
Lutz (1999a)	bivariate	no	yes	yes	yes (as Grilli and Yang)	yes
Newbold *et. al.* (2000)	ARMA	no	possible	possible	yes two series, no index	possible
Newbold and Vougas (1996)	ARMA	no	possible	possible	no	possible
Powell (1991)	Bivariate	no	no	yes	no	yes
Reinhart and Wickham (1994)	ARMA	BNDC and Kalman filter	possible	yes	yes, commodity groups	yes
Sapsford *et. al.* (1992)	ARMA	no	yes	no	no	yes
Sapsford (1985)	As Spraos (1980) + slope and intercept dummies for 1950.	no	yes	not tested	no	yes
Spraos (1980)	Mainly regression on time trend	no	yes	not tested	no	no
von Hagen (1989)	TS vs Bivariate	no	no	yes	yes	yes

BNDC: Beveridge Nelson Decomposition, TS: Trend stationary model

Commodity price series used represented the price of primary commodities relative to the price of manufactured goods in all cases.

Chapter 2

Data Series and Methodology

Chapter 2: Data Series and Methodology

2.1. Description of the Data Series Used

2.1.1. The original data sets

The data used in this study are an extension of the data set used originally by Grilli and Yang (1988) to compile a composite commodity index, which has since become known as the GYCPI[1]. The original Grilli and Yang data covered a total of 24 non-fuel commodity price series for the following commodities: Coffee, Cocoa, Tea, Rice, Wheat, Maize, Sugar, Beef, Lamb, Bananas, Palm Oil, Cotton, Jute, Wool, Hides, Tobacco, Rubber, Timber, Copper, Aluminium, Tin, Silver, Lead and Zinc.

The data set provides annual averages of commodity price data for the period 1900-1986. The data are in US$ and are indexed to the 1977-1979 average of the original current price data. Grilli and Yang used the United Nations Manufacturing Unit Value Index (MUV) and the United States Manufacturing Price Index (USMPI) as alternative deflators for the indexed primary commodity price series. In the present study, only the MUV is used. As a composite index, the MUV-G5 is compiled on the basis of trade flows from the leading five industrial countries to lower income countries[2]:

---

[1] GYCPI stands for Grilli and Yang Commodity Price Index (*cf.* Grilli and Yang (1988)). Grilli and Yang use several versions of the index with different manufactured commodity price series in the denominator and for subgroups of the overall commodity index as well as an index including fuel prices.

[2] The MUV-G5 index described here is alternatively referred to as MUVUN (as in Grilli and Yang) or simply MUV in the literature. Here, MUV will be used, aside from the specific references to the current definition of the MUV-G5 in this chapter.

> "The MUV index is a composite index of prices for manufactured exports from the five major (G-5) industrial countries (France, Germany, Japan, the United Kingdom, and the United States) to low- and middle-income economies, valued in U.S. dollars. The index covers products in Standard International Trade Classification (SITC) groups 5-8. To construct the MUV G-5 index, unit value indexes for each country are combined using weights determined by each country's export share." (Worldbank, World Development Indicators 2000).

The fact that the MUV-G5 is constrained to exports from the leading industrial countries to low and middle income countries, rather than merely relying on price or unit value data for the above mentioned SITC categories, makes the MUV-G5 as a deflator more representative of developing countries' Net Barter Terms of Trade. More specifically, it avoids the inclusion of manufactured products in the later stages of the product life cycle. Some of these products can be more typical of developing countries' economies and their prices may show characteristics more representative of the commonly assumed behaviour of primary commodities than of those associated with manufactured exports from developed economies. (Sarkar (1997) for example works on this assumption, Kaplinsky (1999) provides some empirical support.) (The problems of attempting to link commodity categories to structural features such as comparative advantage or factor endowments, even for very high digit SITC categories, has been extensively documented in the context of the debate on the impact of trade on labour markets. (*cf.* Wood (1994)).

With respect to the purpose of identifying the pressure on developing countries' terms of trade from the development of relative primary commodity prices there is also no reason to expect the MUV index to be inferior to the USMPI. On the contrary, in the case of those developing countries whose trade is not strongly dominated by the USA, the use of the MUV as a deflator should make the relative commodity price series obtained more relevant to developing countries' terms of trade. In the discussion concerning the Prebisch Singer Hypothesis, no case seems to have been made in favour of using the USMPI rather than the MUV and the MUV has been the most frequently used deflator for the GYCPI so far[3].

Data quality and alternatives

Grilli and Yang compile their commodity price index in preference over earlier composite price indices. Among the previously established commodity price indices, the Lewis index[4] does not extend beyond 1938. The Economist commodity price index, by far the longest composite index available, goes back to 1851 and is updated regularly. However, some of the earlier data series are incomplete, it has been subject to various revisions and is based on trade weights of industrialised countries (*cf.* Grilli and Yang (1988)).

While there are reasons to aim for a new commodity index, the Grilli and Yang index has problems of its own. Aside from the required interpolations for the MUV, there are some reasons to doubt the adequacy of part of the data series used. Current Worldbank commodity price data are obtained from a variety of sources, covering not only c.i.f. and f.o.b. prices but also prices recorded at auctions. Clearly, these differences give rise to some concern in itself (*e.g.* regarding the role of transport costs). Yet one advantage of the Worldbank series arises from the fact

---

[3] A detailed listing of data sources used in various studies can be found in Appendix I.ii.

[4] The Lewis index and other early commodity price indices are reviewed in some detail by Spraos (1980), Cashin *et. al.* (2001) cover the Economist Index in some detail.

that the price data are generally averaged over month, years and, where appropriate over auctions.

While averages -as opposed to individual data points- are not ideal from a statistical point of view they tend to be preferred with respect to their economic interpretation if they are taken as representative for one particular period. Average prices, both in the intertemporal sense and in the sense of accounting for fixed factors, should be more representative of prevailing price incentives than individual price quotations which may merely be indicative of the marginal price prevailing for one particular transaction. There is a possibility that this problem may be relevant for the earlier Grilli and Yang data. Grilli and Yang (1988) say little about the origin of their commodity data while Cuddington (1992) describes them as "free market commodity price quotations". This could be a reference to recorded marginal prices.

A further problem is that of the characteristics of the aggregate index and its component parts. The GYCPI, as well as the Lewis index[5] show signs of structural instability around 1920/1921. This is not observed in any of the sub-indices or individual commodity price series. Cuddington (1992) suggests that this may be due to the mixed incidence of integrated and stationary series in the index and quotes further research linking aggregation problems to the use of arithmetic rather than geometric averages in the construction of the index.

These problems then provide further cause to guard against the use of a composite index. In spite of the possible problems with earlier recorded data, it remains one

[5] And the Economist index with respect to its imputed rate of decline (see Cashin *et.al.* (2001)).

of the attractions of the Grilli and Yang data set that a large number of individual price series is available over a long common time horizon.

Further commodity data from the World Bank

Further commodity data for a number of primary commodities were obtained in current dollar terms and deflated by the MUV index. The data series covered the period from 1960-1995. Where possible[6], these data were further extended to cover the years 1996-1998 and to update the data for 1995, where there were discrepancies at the two digit level and where data were available. The main data series used were the primary commodity price data from the World Bank ('pinksheet'). These data were available in current US$ terms and were deflated using the MUV. Price data for Tobacco were obtained from table 6.5 of the World Development Indicators (Primary commodity prices)[7].

Chaining the Series

Both data series overlap for the period 1960-1986. The World Bank data series in current US$ can be converted to fit the CYCPI by applying the following formula:

[2.1.1] $$G_y = \frac{C_i}{Base}100$$

where G_y is the value of the GYCPI in year i, C_i the commodity price in current US$ in year i and *Base* the arithmetic mean of the relevant commodity's price for the years 1977-1979. New data for the MUV, which were indexed to base year

[6] This was not possible for Tea and Tobacco, where data series did not extend beyond 1997 and in the case of Hides were no new data were available after 1995.

[7] Data sources are listed separately in the Bibliography. The Pinksheet gives data for Robusta and Arabica Coffee and in the case of Timber for logs from Malaysia and Cameroon. Data for Coffee were extended from data for Robusta Coffee and in the case of Timber from price data for Malaysian logs. In both cases a closer correspondence was obtained for data during the period 1960-1986 where data were available for the original data set and the new data.

1990, were likewise indexed to their 1977-1979 average. After indexing the commodity price series and the MUV index to their 1977-1979 averages, commodity price series were updated from the more recent world bank data series from 1960 onward, so as to incorporate more recent revisions where discrepancies between the original and more recent data series exist.

This yielded a continuous version of the Grilli and Yang commodity price data relative to the United Nations Manufacturing Unit Value index, and indexed to their 1977-1979 average as in the original study by Grilli and Yang. Finally, natural logarithms were taken of the chained series. (The data series obtained are listed in appendix II.i.).

2.1.2 Description of the Data Series

To gain an initial impression of the data series obtained, some basic numerical measures of concentration and dispersion were calculated. Table 2.1. reports the Arithmetic Mean and Standard Deviation as well as the first and last observation and the minimum and maximum values for each commodity's sample data.

Table 2.1.1. Summary Data for Relative Commodity Price Series

Commodity	Average	Std Dev.	First Obs.	Last Obs.	Min.	Max.
Coffee	-0.794	0.427	-1.157	-0.926	-1.698	0.350
Cocoa	-1.016	0.506	-0.485	-1.319	-1.864	0.222
Tea	0.015	0.387	0.253	-0.672	-1.078	0.700
Rice	0.153	0.431	0.444	-0.583	-0.861	0.915
Wheat	0.279	0.391	0.333	-0.618	-0.618	1.088
Maize	0.304	0.419	0.138	-0.602	-0.743	1.212
Sugar	0.230	0.515	0.821	-0.540	-0.910	1.611
Beef	-0.868	0.680	-1.607	-0.816	-1.906	0.473
Lamb	-0.858	0.697	-1.823	-0.306	-2.075	0.149
Banana	0.217	0.233	-0.116	-0.077	-0.314	0.782
Palm Oil	0.002	0.454	0.126	-0.463	-1.264	1.341
Cotton	0.197	0.405	0.151	-0.689	-0.837	0.888
Jute	0.145	0.435	-0.050	-0.941	-0.942	0.956
Wool	0.495	0.566	0.665	-0.738	-0.869	1.339
Tobacco	-0.251	0.432	-1.162	-0.898	-1.162	0.754
Hides	0.057	0.468	0.217	-0.762	-1.287	1.031
Rubber	0.569	0.922	2.029	-0.931	-0.931	2.629
Timber	-0.427	0.394	-1.070	-0.282	-1.114	0.572
Copper	0.046	0.357	0.397	-0.522	-0.588	0.938
Aluminium	0.419	0.599	1.411	-0.394	-0.592	1.843
Tin	-0.835	0.397	-1.117	-1.441	-1.644	0.045
Silver	-0.749	0.402	-0.517	-0.825	-1.472	0.859
Lead	-0.323	0.331	-0.265	-1.033	-1.319	0.262
Zinc	0.024	0.260	-0.130	-0.117	-0.581	1.012

Data for Tea and Tobacco are for the period 1900-1997 only, Data for Hides were available only until 1995. All other data series cover the period 1900-1998. All Data series give commodity prices relative to MUV in natural logarithms on an annual basis.
Std Dev.: Standard deviation, Obs.: Observation, Max.: Maximum, Min.: Minimum.

These rather crude descriptive statistics can be complemented by graphical illustrations of the time path of relative primary commodity prices. Appendix II.ii. contains Graphs for all the relative commodity price series in the sample, showing their development between 1900 and 1998 (or 1997 for Tea and Tobacco, 1995 for Hides). In the cases of Tea and Tobacco, it was has not been possible to extend the data series beyond 1997, and for Hides no data were available after 1995. Since the series for Hides have not been updated for a prolonged period of time, it is likely that no updated values will be available for the evaluation of forecasts at a later stage. The series is therefore dropped from the study, except for those parts of Chapter four, where its inclusion is of interest in the context of tests for the significance of the trend coefficient.

The time series graphs for some commodities do indeed give the impression of a long run trend, although in some cases, such as Timber, this appears to be positive. The notion of a secular trend does seem plausible for Hides, Rubber, Timber and Aluminium. For other commodities such as Coffee, Cocoa and Zinc there is no obvious evidence of a trend and it appears plausible that these series follow a volatile but stationary time path. In yet other cases, the picture is even less clear. For some series like Tea, Rice Wheat, Maize, Beef, Lamb, Cotton and Palm Oil the presence of an underlying trend as well as stationarity around a structural break seem to be possible scenarios, in yet other cases (Banana, Tobacco, Jute, Wool, Sugar, Tin, Silver, Lead and Copper) the series do not appear stationary but there is no indication of a clear pattern of trends or discrete shifts either.

The one common feature which the time series covered do share is a high degree of volatility over time. Even in those cases where the presence of a trend is strongly suggested by a plot of the data (*e.g.* Timber and Hides) volatility around this trend is a salient feature illustrated by the time series graph. Zinc, the one price series which most clearly appears to be stationary, is clearly subject to extensive fluctuations.

Decisions on the presence of a deterministic trend in the data generating process underlying the series as well as decisions as to how those data series should best be modelled should be taken with care. The questions of the presence of secular trends is clearly important in the context of forecasting as well as in commenting on those issues in trade policy discussions for which the assumption of a secular trend is crucial.

2.2 Basic Econometric Methodology

This section outlines the basic methodology employed in modelling time series and in testing for stationarity in levels allowing for the presence of a deterministic trend (hereafter referred to as trend-stationarity). Deviations from and extensions to this basic methodology will be discussed in later chapters as they are used.

2.2.1. Modelling univariate time series.

In modelling univariate time series in the stationary case, the time series in question can be represented as an ARMA(p,q) process of the form:

[2.2.1] $$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

where p corresponds to the number of autoregressive lags and q to the number of moving average terms, and where

[2.2.2] $$\varepsilon_t \sim \mathrm{N}(0, \sigma_\varepsilon^2)$$

and the definition of stationarity in the sense of 'weak' or covariance stationarity requires that for a stationary series:

1. the mean of the variable under consideration is constant over time,
2. the residual variance is finite and constant over time

3. the covariance between two observations y_t and y_{t-i} depends only on the time lag between those observations not on the point in time when the observations in question occur (*cf.* Granger and Newbold (1986)).

An alternative way of expressing the ARMA(p,q) process identified in equation [2.2.1] is given by:

[2.2.3] $$y_t = u_t,\ u_t - \sum_{i=1}^{p} \phi_i u_{t-i} = \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

so that

[2.2.4] $$y_t = \sum_{i=1}^{p} \phi_i u_{t-i} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Since by [2.2.3] it is generally the case that $y_{t-i} = u_{t-i}$ [2.2.1] can be obtained from [2.2.4] by substitution.

The stationarity assumption is not fulfilled if one or more of the autoregressive parameters take a value of one (*i.e.* if there is a unit root in the characteristic equation (*cf.* Johnston and Dinardo (1997)). The case can be illustrated for the case of a first order autoregressive process:

[2.2.5] $$y_t = \phi y_{t-1} + \varepsilon_t$$

For such an AR(1) process with zero mean[8], the variance is:

[2.2.6] $$V(y_t) = \gamma(0) = E(y_t^2)$$

Since [2.2.3] can be shown, by a process of consecutive substitution, to be equivalent to:

[2.2.5'] $$y_t = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i}$$

[8] The variance expression can be generalised to a case of non zero mean as shown in Johnston Dinardo (1997), p.57*ff.*

[2.2.6] is equivalent to:

[2.2.6']
$$\gamma(0) = E\left(\sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i}\right)^2 = \sum_{i=0}^{\infty} \phi^{2i} E(\varepsilon_{t-i}^2) = \sigma^2 \sum_{i=0}^{\infty} \phi^{2i}$$

this can in turn be written as:

[2.2.7]
$$\gamma(0) = \frac{\sigma^2}{(1-\phi^2)}$$

For the case where $\phi = 1$ one may then consider a time series for a finite period of T observations with y_0=0, so that $\Sigma_{i=0}^{T} \phi^{2i}$ =T and [2.2.6'] becomes:

[2.2.8]
$$V(y_t) = \gamma(0) = \mathrm{T}\sigma^2$$

while [2.2.7] is not defined for this case. The variance expression for $\phi = 1$ is clearly dependent on time and thus violates the above stationarity assumptions. Equation [2.2.5.] would here express the pure random walk case if $\phi = 1$:

[2.2.9]
$$y_t = y_{t-1} + \varepsilon_t$$

or, rearranging

[2.2.9']
$$\Delta y_t = (1-\mathrm{L})y_t = \varepsilon_t$$

One should thus consider the possibility that a series which is not stationary in levels may be stationary in first differences, *i.e.* integrated of order one (I(1)). For the more general ARIMA(p,1,q) case this would be:

[2.210]
$$\Delta y_t = v_t,\ v_t - \sum_{i=1}^{p} \phi_i v_{t-i} = \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

which, substituting for v_t and rearranging can be expressed as:

[2.2.11]
$$\Delta y_t = \sum_{i=1}^{p} \phi_i v_{t-i} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

and, with $\Delta y_{t-i} = v_{t-i}$ this can be expressed as:

$$[2.2.12] \qquad \Delta y_t = \sum_{i=1}^{p} \phi_i \Delta y_{t-i} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Considering a data generating process characterised by a constant and deterministic linear trend, the simplest case is:

$$[2.2.13] \qquad y_t = \alpha + \beta t + \varepsilon_t,\ \varepsilon_t \sim \mathrm{N}(0, \sigma_\varepsilon^2)$$

The equivalent for an I(1) data generating process with trend then corresponds to the first differenced version of [2.2.3], *i.e.* random walk plus drift case, which is:

$$[2.2.14] \qquad (1 - \mathrm{L})y_t = \beta + \varepsilon_t$$

since obviously a first differenced constant is eliminated from the expression and a first differenced linear trend reduces to a constant.

It has been illustrated, for the AR(1) case, in equations [2.2.5'] to [2.2.8] that a pure autoregressive process can be inverted to yield a moving average process regardless of stationarity characteristics. It is not the case, however, that a pure moving average process is always invertible. More specifically, this will not be the case if in an equation like:

$$[2.2.15] \qquad y_t = \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

One or more of the θ_i parameters take a value of one, *i.e.* if at least one of the moving average parameters has a unit root. Considering the MA(1) process

$$[2.2.16.] \qquad y_t = \theta \varepsilon_{t-1} + \varepsilon_t$$

so that

$$[2.2.17] \qquad y_{t-1} = \theta \varepsilon_{t-2} + \varepsilon_{t-1}, \text{ and}$$

$$[2.2.18] \qquad \theta y_{t-1} = \theta^2 \varepsilon_{t-2} + \theta \varepsilon_{t-1}$$

This can be solved for $\theta\varepsilon_{t-1}$ and substituted back into [2.2.16]. Repeated substitution will then yield:

[2.2.19] $$y_t = -\sum_{i=1}^{\infty}(-\theta)^i\varepsilon_{t-i} + \varepsilon_t$$

clearly, where the value of the moving average parameter is within the range $0 \leq \theta < 1$ the autoregressive coefficients obtained by inversion will approach zero for larger lags. For the case of $\theta = 1$, however, the autoregressive terms will not converge on zero as the number of lags goes to infinity. A Moving average process is therefore not invertible if at least one of the moving average parameters takes a value of one or above.

2.2.2. Testing for unit roots

The question of the order of integration of a time series is of crucial importance when considering the question of a deterministic trend or drift term and when attempting to assess the statistical significance of the estimated trend or drift coefficient. It is a well documented result (*cf.* Newbold and Granger (1974)) that ordinary t-tests reject the null hypothesis of a zero trend coefficient too frequently when the data generating process[9] is I(1) and a trend stationary model for an I(0) process is fitted.

It is therefore common practise to pre-test the data series in question for the presence of unit roots in the autoregressive component to decide on the order of integration. In the present case of relative primary commodity prices in logarithms one can expect the series to be integrated of either order 0 or 1. The most common

[9] Throughout, I(1) is used to refer to difference stationary processes, I(0) to refer to data series that are stationary in levels, possibly around a trend.

way of testing for the presence of unit roots is the application of the Augmented Dickey Fuller (ADF) test, in this case including a constant and trend, to the data in levels. The ADF testing equation allowing for a linear trend and constant takes the form:

[2.2.20] $$\Delta y_t = a + \beta t + \rho y_{t-1} + \sum_{i=1}^{n} \gamma_i \Delta y_{t-i} + \varepsilon_t$$

Where the coefficients for the trend and constant are as before, ρ is the coefficient on the lagged dependent variable and the γ_i are the coefficients on the lagged differenced terms. The number of lagged differenced terms should be large enough to avoid the occurrence of serial correlation in the error term. The true number n of autoregressive lags can generally be identified if the initial testing equation has a number of $m>n$ lags and insignificant lagged terms are then successively eliminated until either the last term remains significant, or the testing equation reduces to an ordinary Dickey-Fuller test including a trend and constant.

The null hypothesis of a unit root would then be tested by evaluating the t-ratio on the estimate of the autoregressive coefficient $\hat{\rho}$, and testing $H_0 : \rho = 0$ using a set of critical values provided by Dickey and Fuller (these values are quoted *e.g.* in Enders (1995) and Johnston and Dinardo (1997)).

To test for the presence of unit roots in a first differenced data series allowing for a trend, the ADF testing equation now takes the form:

[2.2.21] $$\Delta^2 y_t = \beta + \rho \Delta y_{t-1} + \sum_{i=1}^{n} \gamma_i \Delta^2 y_{t-i} + \varepsilon_t$$

where a constant is included to allow for a linear trend. One would then proceed as above in determining the number of autoregressive lags and evaluating the significance of $\hat{\rho}$ using the appropriate set of critical values, which is different from the one for [2.2.20] (*cf.* again Johnston and Dinardo (*op.cit.*)).

Since unit root tests are commonly used to make *a priori* decisions on stationarity and the order of integration of a time series for subsequent analysis, and since the results obtained subsequently depend crucially on the inference on unit roots made previously (*cf.* Newbold and Vougas (1996) and Newbold *et. al.* (2000)), the issue of *a priori* testing should not be taken lightly. One fundamental problem with Dickey-Fuller type tests -as with significance tests in general- is the essentially arbitrary nature of the significance level chosen. Indeed in the case of testing for unit roots -or, by implication, for stationarity- it is not at all clear in principle, how the null and alternative hypotheses should be specified. In the case of stationarity tests, such as the one developed by Leybourne and McCabe (1994), Leybourne and McCabe (1999) and the one developed by Kwiatkowski *et. al.* (1992) the null and alternative hypotheses are effectively reversed when compared to unit root tests.

There are further problems with the application of unit root tests. Perron (1989) pointed out that unit root tests have low power if a structural break is not accounted for in an otherwise stationary series. For this case, Perron proposes the inclusion of dummy variables to account for the presence of a structural break. This procedure is well established for the case of a single exogenously determined structural break. There does not appear to be a generally accepted method for dealing with multiple or endogenously inferred structural breaks though.

Finally, it has been shown by Agiakloglou and Newbold (1992) that rejections of the null hypothesis of a unit root occur too often if a moving average coefficient close to one is present in the data generating process. This problem can be dealt with in the test by Leybourne and McCabe (1994) which is also reported to be robust against overfitting in terms of autoregressive terms in the testing equation. (*Cf.* Leybourne and McCabe (1994), the Leybourne McCabe test is discussed in detail when it is used in Chapter 4.)

One approach (adopted *e.g.* by Ahrens and Sharma (1997) and León and Soto (1997)) is to anticipate the problems mentioned above as far as possible and to select the most appropriate unit root test for *a priori* testing. The selection of the most appropriate pre-test is itself not always unproblematic though and any mistaken inference at the pre-test stage can be expected to influence the later analysis. Given the overall uncertainty surrounding *a priori* testing and the possibly far reaching consequences of mistaken inferences at this stage, the present study will give extensive consideration to models based on the assumption that the data generating process is stationary as well to those that assume an I(1) process.

2.2.3. Model estimation and selection

The ARIMA(p,d,q) models with d=0,1 for all commodity price series were estimated using exact Maximum Likelihood estimation, as implemented in the time series package for GAUSS. The general model specification adopted for the model in levels was:

[2.2.22] $$y_t = a + \beta t + u_t,\ u_t - \sum_{i=1}^{p} \phi_i u_{t-i} = \varepsilon_t - \sum_{j=0}^{q} \theta_j \varepsilon_{t-j}$$

where the residual, u_t, follows an ARMA(p,q) process and a trend and constant term are included in the main equation. The equivalence of pure ARMA processes with the autoregressive terms expressed in terms of either the residual or the dependent variable was demonstrated above in equations [2.2.1] to [2.2.4]. One should note here that no such equivalence holds for the values of the constant and trend coefficients. It is straightforward to show that for an ARMA(p,q) model of form [2.2.1] including constant and trend,

[2.2.1']
$$y_t = a + bt + \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

The constant and trend coefficients would correspond to those in [2.2.22] as in $b = (1 - \Sigma_{i=1}^{p} \phi_i)\beta$ and $a = (1 - \Sigma_{i=1}^{p} \phi_i)\alpha + \beta \Sigma_{i=1}^{p} i\phi_i$. For the difference stationary model, the basic specification adopted was:

[2.2.23]
$$\Delta y_t = \beta + v_t,\ v_t - \sum_{i=1}^{p} \phi_i v_{t-i} = \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

Where the residual process for the series in first differences is again modelled as an ARMA(p,q) process. If the difference stationary model were to be represented as in [2.2.12] but including a constant term, *i.e.*

[2.2.12']
$$\Delta y_t = b + \sum_{i=1}^{p} \phi_i \Delta y_{t-i} - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

it would again be the case that $b = (1 - \Sigma_{i=1}^{p} \phi_i)\beta$ as above. The optimal representation of the time series was then selected as follows: for all I(0) and I(1) series, all possible ARIMA(p,d,q) specifications with $p + q \leq 5$ were estimated. The optimal model specification was then inferred from the lowest value for the Schwarz Bayesian Information Criterion (SBC).

In principle, the appropriate ARIMA model specification can be selected by, at least, either the Akaike Information Criterion or the SBC. Harvey (1993) defines the AIC as:

[2.2.24] $\text{AIC} = -2\ln L(\psi) + 2n,$

where $\ln L(\psi)$ is the value of the maximised likelihood function and n the number of estimated coefficients in the estimating equation. The SBC is defined as:

[2.2.25] $\text{SBC} = -2\ln L(\psi) + n\ln(\text{T})$

where T is the number of useable observations in the sample and the remaining components of the equation are as in [2.2.24]. In either expression, inclusion of the negative of the maximised log likelihood function assures that the value of the test statistic tends to be lower for model specifications with a closer fit. The $2n$ and $n\ln(\text{T})$ terms respectively result in the test statistic tending towards lower values for more parsimonious model specifications. Both test statistics are designed to select ARIMA models on the basis of a trade-off between goodness of fit and parsimonious specification. The AIC, however, is known to consistently select overparameterized models (*cf.* Harvey *op. cit.*) so that in the present study the SBC is used as the selection criterion.

2.2.4. Further considerations for the formulation of forecast models

It has been pointed out by Granger and Newbold (1986) that optimal predictions from a univariate ARIMA(p,d,q) model can be obtained if the correct model corresponding to the data generating process has been identified. It should be clear from the preceding discussion that such a correct model identification can not be taken for granted. Uncertainty about the accuracy of forecasts therefore arises from uncertainty about model selection as well as from the standard errors obtained with the coefficient estimates in the selected forecast model. No attempt will be undertaken here to quantify the uncertainty associated with the question of appropriate model selection -once a forecast model has been decided upon. The influence of standard errors for the estimated coefficients in the forecast model can be incorporated through confidence intervals for the forecasts obtained. It is worth bearing in mind that the imputed order of integration of the data generating process has an immediate impact here. The confidence intervals for forecasts from a difference stationary model will widen for longer forecast horizons, while the confidence intervals for forecasts from trend stationary models will not. In the context of uncertainty about the order of integration, this raises the additional issue of the lowest cost of misspecification. (This lower cost tends be attributed to the difference stationary alternative. Chapter 5, contains a more detailed treatment of this issue.)

It has been repeatedly pointed out that the incorporation of trend or drift terms in the estimated ARIMA models can be decisive for the values of the forecasts obtained. In data series which are characterised by substantial volatility as well as the presence of trend or drift components the importance of the trend term or even the desirability of its inclusion can depend on its magnitude. This question will be taken up in Chapter 5, where the impact of correctly including trend or drift terms of different magnitude is reviewed in more depth.

Finally, it should be desirable to quantify the persistence of shocks to the data series. In so far as forecast models are entirely or mainly characterised by their ARMA parameterisation, such a measure of shock persistence should give a basis for assessing the importance of predicted mean reverting movements. While a distinction of trend and cyclical components is straightforward for stationary or trend stationary series, this distinction is more complex in the case of difference stationary data series. Chapter 6 will therefore present computations of the Beveridge-Nelson trend cycle decomposition for those data series where a difference stationary forecast model has been considered.

2.3 Conclusion

In this study, the individual commodity price series underlying the Grilli and Yang commodity price index, deflated by the MUV Manufacturing Unit Value index will be used to evaluate the presence of a trend term in long run primary commodity price series and to obtain forecasts for primary commodity prices. For this purpose the series have been updated to cover the period 1900-1998 where possible.

The estimates will be based on univariate ARIMA models allowing for the presence of a trend or drift term. In contrast to most previous studies on the topic the problems surrounding unit root pre-testing motivate the consideration of stationary and difference stationary model alternatives for the evaluation of possible trend components in the series. Forecast models will be selected taking these findings into account, and trend cycle decompositions will be computed to assess the incidence of permanent shock components.

Appendix II.i. Average Annual Commodity Price Series Deflated by MUV.

This appendix lists the annual average price series for relative primary commodity prices deflated by the Manufacturing Unit Value Index (MUV) and indexing the data series to their 1977-1979 average as described in Chapter 2. (*n.a.* indicates where no data are available.)

Year	Coffee	Cocoa	Tea*	Rice	Wheat
1900	0.315	0.616	1.288	1.559	1.395
1901	0.260	0.610	1.298	1.477	1.410
1902	0.226	0.631	1.414	1.468	1.659
1903	0.230	0.635	1.077	1.857	1.659
1904	0.312	0.633	1.257	1.506	1.879
1905	0.332	0.622	1.221	1.695	1.839
1906	0.303	0.605	1.016	1.890	1.473
1907	0.234	0.697	0.946	1.930	1.622
1908	0.323	0.523	0.945	2.160	2.069
1909	0.342	0.448	1.034	1.700	2.168
1910	0.404	0.421	1.065	1.697	1.910
1911	0.548	0.444	1.074	2.221	1.890
1912	0.606	0.458	1.059	2.497	1.899
1913	0.493	0.509	1.072	1.955	1.550
1914	0.459	0.479	1.130	1.942	2.043
1915	0.373	0.628	1.065	1.915	2.526
1916	0.333	0.432	0.861	1.429	2.219
1917	0.269	0.286	0.921	1.035	2.969
1918	0.276	0.271	0.887	0.971	2.423
1919	0.509	0.369	0.804	0.985	2.289
1920	0.367	0.252	0.603	1.083	2.287
1921	0.236	0.171	0.620	1.469	1.698
1922	0.364	0.227	0.974	1.821	1.590
1923	0.377	0.187	1.245	1.729	1.381
1924	0.543	0.187	1.273	1.880	1.642
1925	0.616	0.230	1.244	1.863	2.101
1926	0.588	0.293	1.414	2.111	2.025
1927	0.521	0.426	1.471	2.066	2.111
1928	0.647	0.345	1.295	1.869	1.925
1929	0.640	0.291	1.311	1.938	1.986
1930	0.381	0.234	1.248	1.558	1.436

1931	0.317	0.181	1.139	1.066	1.051
1932	0.461	0.185	0.818	1.130	1.067
1933	0.354	0.165	1.101	0.859	1.114
1934	0.365	0.165	1.258	0.944	1.277
1935	0.299	0.162	1.213	1.261	1.460
1936	0.319	0.221	1.250	1.238	1.615
1937	0.361	0.267	1.410	1.284	2.251
1938	0.242	0.158	1.265	1.128	1.608
1939	0.254	0.160	1.202	1.193	1.090
1940	0.223	0.155	1.123	1.312	1.061
1941	0.334	0.217	1.351	1.378	0.998
1942	0.341	0.219	1.518	1.305	0.964
1943	0.305	0.196	1.218	1.217	1.407
1944	0.268	0.172	1.045	1.069	1.318
1945	0.264	0.167	1.018	1.041	1.382
1946	0.355	0.213	0.952	1.185	2.238
1947	0.424	0.536	1.092	1.571	2.154
1948	0.422	0.597	1.137	1.630	1.894
1949	0.543	0.347	1.292	1.502	1.681
1950	0.852	0.566	1.265	1.392	1.636
1951	0.849	0.528	1.339	1.241	1.598
1952	0.803	0.516	1.094	1.316	1.642
1953	0.837	0.564	1.365	1.533	1.551
1954	1.156	0.898	2.014	1.416	1.379
1955	0.891	0.576	1.890	1.255	1.317
1956	0.982	0.402	1.752	1.164	1.263
1957	0.958	0.446	1.591	1.155	1.213
1958	0.712	0.653	1.664	1.210	1.204
1959	0.612	0.539	1.652	1.124	1.247
1960	0.617	0.472	1.699	0.990	1.245
1961	0.588	0.382	1.598	1.079	1.237
1962	0.536	0.354	1.585	1.200	1.332
1963	0.528	0.435	1.532	1.139	1.361
1964	0.651	0.391	1.529	1.070	1.402
1965	0.642	0.281	1.480	1.050	1.224
1966	0.574	0.384	1.391	1.240	1.252
1967	0.528	0.438	1.387	1.599	1.292
1968	0.535	0.534	1.154	1.560	1.246
1969	0.514	0.635	1.019	1.362	1.100
1970	0.632	0.446	1.081	0.961	0.973
1971	0.519	0.338	0.986	0.805	1.038
1972	0.529	0.370	0.903	0.857	1.077
1973	0.569	0.562	0.784	1.676	1.862
1974	0.494	0.636	0.853	2.428	1.964

1975	0.440	0.457	0.758	1.442	1.465
1976	0.949	0.740	0.829	0.977	1.289
1977	1.419	1.248	1.320	0.956	0.911
1978	0.857	0.974	0.934	1.139	0.980
1979	0.805	0.832	0.812	0.911	1.086
1980	0.666	0.600	0.767	1.089	1.067
1981	0.549	0.477	0.690	1.213	1.076
1982	0.600	0.405	0.672	0.731	1.002
1983	0.579	0.506	0.828	0.705	1.006
1984	0.648	0.584	1.257	0.652	0.995
1985	0.651	0.545	0.715	0.548	0.880
1986	0.734	0.424	0.590	0.440	0.631
1987	0.390	0.372	0.476	0.461	0.565
1988	0.440	0.276	0.464	0.556	0.677
1989	0.349	0.217	0.528	0.603	0.795
1990	0.273	0.210	0.503	0.517	0.602
1991	0.253	0.194	0.446	0.548	0.559
1992	0.183	0.171	0.464	0.480	0.630
1993	0.203	0.174	0.434	0.423	0.586
1994	0.415	0.210	0.411	0.464	0.604
1995	0.386	0.199	0.340	0.513	0.659
1996	0.327	0.212	0.385	0.568	0.809
1997	0.534	0.249	0.511	0.537	0.656
1998	0.396	0.267	*n.a.*	0.558	0.539

* Data for Tea only until 1997

Year	Maize	Sugar	Beef	Lamb	Banana
1900	1.148	2.272	0.201	0.162	0.891
1901	1.549	1.938	0.211	0.170	0.972
1902	2.049	1.592	0.217	0.175	1.034
1903	1.461	1.705	0.207	0.175	1.070
1904	1.525	2.209	0.211	0.170	1.078
1905	1.612	2.361	0.211	0.170	1.116
1906	1.346	1.702	0.201	0.162	1.096
1907	1.450	1.703	0.186	0.150	1.080
1908	2.041	2.110	0.206	0.159	1.206
1909	1.991	2.077	0.206	0.166	1.167
1910	1.673	2.241	0.205	0.165	1.195
1911	1.791	2.489	0.204	0.165	1.272
1912	1.962	2.096	0.200	0.161	1.293
1913	1.779	1.566	0.197	0.159	1.276
1914	2.127	2.234	0.203	0.164	1.369
1915	2.172	2.728	0.424	0.366	1.300

1916	1.978	2.911	0.373	0.307	1.091
1917	3.361	2.583	0.312	0.285	1.052
1918	2.688	1.952	0.345	0.297	1.048
1919	2.510	2.201	0.404	0.336	0.893
1920	2.092	4.860	0.244	0.243	0.939
1921	0.986	1.493	0.223	0.218	1.024
1922	1.210	1.512	0.149	0.129	1.090
1923	1.599	2.716	0.161	0.203	1.153
1924	1.875	2.062	0.162	0.255	1.196
1925	1.981	1.189	0.200	0.293	1.350
1926	1.514	1.241	0.201	0.205	1.497
1927	1.846	1.560	0.192	0.226	1.578
1928	2.088	1.288	0.237	0.278	1.545
1929	2.074	1.056	0.261	0.265	1.616
1930	1.860	0.770	0.240	0.242	1.651
1931	1.430	0.848	0.488	0.495	1.913
1932	1.017	0.654	0.428	0.458	2.187
1933	1.188	0.799	0.358	0.466	1.998
1934	1.635	0.828	0.244	0.414	1.672
1935	2.095	1.124	0.265	0.696	1.724
1936	2.154	1.231	0.273	0.799	1.664
1937	2.595	1.225	0.381	0.861	1.562
1938	1.315	0.966	0.323	0.712	1.540
1939	1.318	1.100	0.337	0.638	1.762
1940	1.390	0.906	0.389	0.596	1.760
1941	1.599	1.058	0.427	0.578	1.713
1942	1.628	1.366	0.475	0.517	1.527
1943	1.802	1.175	0.449	0.494	1.424
1944	1.736	1.045	0.388	0.443	1.337
1945	1.741	1.211	0.378	0.417	1.358
1946	2.403	1.444	0.429	0.446	1.523
1947	2.513	1.616	0.343	0.351	1.321
1948	1.145	1.394	0.278	0.249	1.320
1949	1.823	1.464	0.235	0.234	1.562
1950	2.162	1.925	0.178	0.135	1.793
1951	1.929	1.849	0.150	0.133	1.513
1952	1.643	1.332	0.156	0.126	1.502
1953	1.647	1.136	0.183	0.173	1.566
1954	1.630	1.110	0.180	0.147	1.644
1955	1.349	1.091	0.210	0.179	1.605
1956	1.368	1.123	0.348	0.322	1.559
1957	1.249	1.649	0.304	0.307	1.624
1958	1.262	1.130	0.329	0.310	1.517
1959	1.222	0.959	0.751	0.506	1.354

1960	1.168	1.031	0.948	0.868	1.356
1961	1.217	0.939	0.863	0.765	1.293
1962	1.336	0.943	0.886	0.823	1.207
1963	1.450	2.742	0.844	0.855	1.559
1964	1.453	1.861	1.045	0.952	1.552
1965	1.422	0.666	1.088	0.970	1.440
1966	1.484	0.565	1.225	0.897	1.353
1967	1.232	0.609	1.227	0.849	1.375
1968	1.224	0.602	1.291	0.820	1.331
1969	1.275	0.970	1.381	0.896	1.318
1970	1.300	1.016	1.385	0.829	1.292
1971	1.233	1.162	1.356	0.839	1.046
1972	1.085	1.753	1.369	1.037	1.094
1973	1.639	1.960	1.605	1.161	0.968
1974	1.812	5.007	1.036	0.975	0.887
1975	1.477	3.081	0.782	0.910	1.069
1976	1.369	1.716	0.919	0.952	1.100
1977	1.057	1.095	0.797	0.940	1.069
1978	0.970	0.917	0.983	1.045	0.972
1979	0.983	1.000	1.171	1.006	0.972
1980	0.972	2.705	1.022	1.116	1.032
1981	1.011	1.587	0.912	1.055	1.088
1982	0.858	0.804	0.895	0.930	1.031
1983	1.092	0.827	0.935	0.773	1.208
1984	1.115	0.520	0.890	0.786	1.064
1985	0.913	0.402	0.836	0.749	1.086
1986	0.605	0.508	0.689	0.741	0.925
1987	0.476	0.517	0.716	0.678	0.867
1988	0.626	0.727	0.704	0.704	0.983
1989	0.658	0.918	0.723	0.683	1.132
1990	0.610	0.853	0.683	0.740	1.060
1991	0.587	0.596	0.694	0.635	1.072
1992	0.546	0.577	0.613	0.682	0.869
1993	0.536	0.638	0.656	0.761	0.816
1994	0.545	0.747	0.563	0.752	0.781
1995	0.577	0.756	0.426	0.611	0.730
1996	0.812	0.713	0.417	0.805	0.807
1997	0.605	0.715	0.458	0.875	0.912
1998	0.548	0.583	0.442	0.736	0.926

Year	Palm Oil	Cotton	Jute	Wool	Hides*
1900	1.134	1.163	0.952	1.945	1.243
1901	1.133	1.233	0.899	1.721	1.263
1902	1.229	1.309	0.888	1.919	1.305
1903	1.250	1.596	0.972	2.069	1.243
1904	1.196	1.492	0.982	2.160	1.320
1905	1.175	1.390	1.296	2.636	1.643
1906	1.254	1.454	1.560	2.466	1.766
1907	1.287	1.401	1.328	2.312	1.325
1908	1.164	1.446	1.058	2.139	1.348
1909	1.225	1.652	0.889	2.495	1.811
1910	1.491	1.950	1.010	2.317	1.593
1911	1.471	1.698	1.385	2.032	1.648
1912	1.373	1.499	1.397	2.223	1.963
1913	1.473	1.648	1.763	1.945	2.055
1914	1.699	1.505	1.924	2.160	2.417
1915	1.736	1.386	1.407	2.531	2.805
1916	2.015	1.679	1.675	2.421	2.458
1917	2.481	2.201	1.786	3.798	2.452
1918	3.824	2.246	1.467	3.626	1.551
1919	1.941	2.356	1.641	3.349	2.550
1920	1.270	1.735	1.136	2.815	1.762
1921	0.836	1.349	0.871	1.772	0.812
1922	0.990	1.908	1.253	2.919	1.213
1923	1.017	2.429	1.092	3.293	1.035
1924	1.097	2.360	1.290	3.293	0.983
1925	1.223	1.875	2.167	3.191	1.150
1926	1.192	1.535	2.029	2.806	1.084
1927	1.172	1.622	1.603	2.812	1.638
1928	1.186	1.826	1.653	2.965	1.981
1929	1.248	1.705	1.623	2.603	1.424
1930	1.009	1.281	1.036	2.059	2.153
1931	0.908	0.953	0.937	2.081	0.961
1932	0.868	0.971	0.884	1.873	0.767
1933	0.776	1.194	0.885	2.388	1.148
1934	0.932	1.299	0.841	2.468	0.904
1935	1.358	1.375	1.003	2.309	1.055
1936	1.376	1.392	1.060	2.833	1.209
1937	1.483	1.199	1.153	3.070	1.537
1938	1.123	0.933	0.981	2.017	1.027
1939	1.264	1.090	1.462	2.615	1.306
1940	1.189	1.110	1.270	2.767	1.254
1941	1.506	1.447	1.134	2.953	1.373

1942	1.592	1.612	0.918	2.780	1.240
1943	1.027	1.505	1.093	2.459	1.107
1944	0.902	1.369	1.169	2.179	0.972
1945	0.878	1.492	1.110	2.103	0.946
1946	0.867	1.889	1.372	1.812	1.103
1947	1.485	1.781	1.847	1.809	1.437
1948	1.781	1.656	2.190	2.347	1.363
1949	1.404	1.708	1.701	2.533	1.293
1950	1.399	2.231	2.070	3.332	1.644
1951	1.868	2.095	2.600	3.817	1.648
1952	1.093	1.824	1.526	2.285	0.843
1953	1.007	1.730	1.263	2.493	0.869
1954	1.055	1.753	1.516	2.512	0.716
1955	1.085	1.737	1.359	2.070	0.679
1956	1.200	1.614	1.369	1.915	0.756
1957	1.204	1.556	1.621	2.230	0.734
1958	1.152	1.593	1.432	1.655	0.765
1959	1.168	1.543	1.412	1.698	1.187
1960	1.075	1.134	2.529	1.399	0.910
1961	1.074	1.151	1.679	1.365	1.051
1962	0.982	1.105	1.351	1.329	1.125
1963	1.029	1.097	1.336	1.594	1.374
1964	1.089	1.082	1.636	1.633	1.504
1965	1.230	1.047	1.849	1.356	1.238
1966	1.028	0.979	2.152	1.444	1.277
1967	0.964	1.038	2.017	1.241	0.973
1968	0.735	1.084	1.903	1.304	0.846
1969	0.749	0.930	1.905	1.017	0.965
1970	1.011	0.910	1.717	0.862	0.742
1971	0.963	1.012	1.701	0.768	0.553
1972	0.735	0.994	1.632	1.232	0.745
1973	1.103	1.466	1.361	2.274	1.761
1974	1.604	1.257	1.365	1.334	1.231
1975	0.936	0.928	1.290	0.896	0.729
1976	0.865	1.333	1.014	1.099	0.927
1977	1.026	1.115	1.002	1.050	1.004
1978	1.010	0.980	1.081	0.964	0.956
1979	0.972	0.930	0.927	0.994	1.036
1980	0.790	1.027	0.673	0.945	0.882
1981	0.770	0.922	0.600	0.874	0.631
1982	0.610	0.811	0.631	0.815	0.609
1983	0.703	0.963	0.683	0.773	0.526
1984	1.044	0.947	1.226	0.797	0.573
1985	0.712	0.694	1.336	0.766	0.515

1986	0.310	0.472	0.525	0.604	0.539
1987	0.376	0.670	0.571	0.751	0.729
1988	0.447	0.530	0.610	0.899	0.832
1989	0.361	0.639	0.620	0.835	0.635
1990	0.283	0.657	0.642	0.601	0.448
1991	0.323	0.593	0.585	0.511	0.332
1992	0.360	0.433	0.471	0.545	0.363
1993	0.346	0.435	0.404	0.419	0.276
1994	0.467	0.578	0.425	0.522	0.369
1995	0.513	0.644	0.485	0.604	0.467
1996	0.454	0.562	0.631	0.540	*n.a.*
1997	0.493	0.585	0.443	0.589	*n.a.*
1998	0.629	0.502	0.390	0.478	*n.a.*

*For Hides GYCPI data extend until 1960 since current Worldbank data are only available from 1961. The data series for Hides only extends to 1995.

Year	Tobacco*	Rubber	Timber	Copper	Aluminium
1900	0.313	7.605	0.343	1.487	4.100
1901	0.335	7.059	0.336	1.560	4.361
1902	0.363	7.037	0.338	1.098	4.483
1903	0.384	8.971	0.357	1.318	4.483
1904	0.362	9.792	0.329	1.241	4.626
1905	0.349	10.468	0.328	1.510	4.626
1906	0.347	10.324	0.335	1.771	4.489
1907	0.342	9.063	0.331	1.748	5.367
1908	0.391	7.896	0.333	1.245	3.693
1909	0.406	11.363	0.339	1.224	2.831
1910	0.399	13.860	0.358	1.201	2.870
1911	0.411	11.179	0.358	1.167	2.587
1912	0.414	10.553	0.367	1.501	2.758
1913	0.428	8.585	0.385	1.403	2.959
1914	0.460	6.786	0.421	1.317	2.458
1915	0.469	6.693	0.578	1.629	4.388
1916	0.396	6.412	0.738	2.073	6.315
1917	0.462	5.229	0.875	1.739	4.471
1918	0.631	3.377	0.931	1.298	2.416
1919	0.665	2.857	0.701	0.930	2.180
1920	0.958	2.846	0.611	0.812	1.943
1921	0.863	1.288	0.453	0.689	1.595
1922	0.823	1.531	0.440	0.826	1.577
1923	0.778	2.609	0.507	0.891	2.142
1924	0.724	2.300	0.455	0.804	2.276
1925	0.781	6.290	0.493	0.853	2.254

1926	0.720	4.514	0.455	0.883	2.358
1927	0.730	3.653	0.483	0.873	2.344
1928	0.712	2.157	0.500	0.975	2.205
1929	0.727	2.056	0.503	1.272	2.292
1930	0.718	1.049	0.487	0.930	2.328
1931	0.750	0.765	0.387	0.709	2.779
1932	0.694	0.517	0.379	0.586	3.351
1933	0.724	0.795	0.412	0.662	2.998
1934	0.934	1.473	0.436	0.671	2.347
1935	1.127	1.434	0.400	0.704	2.278
1936	1.078	1.915	0.436	0.771	2.278
1937	1.004	2.193	0.555	1.049	2.184
1938	0.983	1.587	0.481	0.762	2.081
1939	0.772	2.080	0.542	0.913	2.274
1940	0.608	2.168	0.714	0.861	1.946
1941	0.697	2.264	0.847	0.845	1.614
1942	0.697	1.986	0.805	0.728	1.265
1943	0.938	1.772	0.748	0.649	1.128
1944	0.993	1.557	0.731	0.570	0.991
1945	0.943	1.516	0.611	0.555	0.965
1946	0.998	1.496	0.663	0.643	0.889
1947	0.830	1.155	0.614	0.807	0.736
1948	0.765	1.186	0.733	0.831	0.808
1949	0.810	1.010	0.546	0.773	0.934
1950	0.924	2.598	0.644	0.939	1.069
1951	0.921	3.151	0.782	0.903	0.968
1952	0.901	2.016	0.800	0.885	0.968
1953	0.991	1.320	0.717	1.098	1.087
1954	1.032	1.316	0.710	1.156	1.159
1955	1.006	2.155	0.783	1.444	1.246
1956	0.956	1.804	0.764	1.545	1.210
1957	1.034	1.628	0.764	1.081	1.267
1958	1.057	1.482	0.700	0.951	1.250
1959	1.066	1.930	0.647	1.152	1.245
1960	2.125	2.140	0.745	1.220	1.255
1961	1.881	1.622	0.759	1.122	1.212
1962	1.657	1.491	0.819	1.119	1.113
1963	1.615	1.406	0.814	1.144	1.071
1964	1.759	1.300	0.682	1.685	1.103
1965	1.484	1.328	0.784	2.229	1.134
1966	1.792	1.198	0.796	2.554	1.095
1967	1.456	0.980	0.840	1.878	1.105
1968	1.362	0.969	0.864	2.068	1.142
1969	1.194	1.208	0.796	2.317	1.153

1970	1.087	0.922	0.831	2.102	1.145
1971	0.999	0.714	0.791	1.524	1.096
1972	0.944	0.654	0.690	1.387	0.916
1973	0.856	1.154	0.985	1.996	0.911
1974	0.923	1.050	0.970	1.889	0.856
1975	1.032	0.705	0.721	1.021	0.910
1976	1.208	0.959	0.970	1.140	1.009
1977	1.134	0.919	0.888	0.971	1.077
1978	0.991	0.966	0.812	0.880	0.970
1979	0.905	1.092	1.252	1.128	0.968
1980	0.801	1.124	1.312	1.131	1.045
1981	0.814	0.882	1.033	0.899	0.902
1982	0.913	0.684	0.994	0.776	0.720
1983	0.968	0.869	0.957	0.854	1.069
1984	1.037	0.799	1.113	0.755	0.949
1985	0.964	0.628	0.857	0.771	0.784
1986	0.833	0.566	0.828	0.634	0.734
1987	0.783	0.629	1.097	0.749	0.910
1988	0.656	0.706	1.018	1.018	1.382
1989	0.848	0.582	0.973	1.123	1.064
1990	0.859	0.491	0.856	0.993	0.846
1991	0.867	0.459	0.905	0.853	0.658
1992	0.817	0.459	0.949	0.798	0.607
1993	0.642	0.444	1.771	0.671	0.553
1994	0.608	0.580	1.348	0.781	0.692
1995	0.558	0.752	1.034	0.917	0.781
1996	0.386	0.694	1.068	0.751	0.682
1997	0.407	0.535	1.066	0.786	0.765
1998	*n.a.*	0.394	0.755	0.593	0.674

* Data for Tobacco are available until 1997 only.

Year	Tin	Silver	Lead	Zinc
1900	0.327	0.597	0.767	0.878
1901	0.193	0.604	0.801	0.858
1902	0.318	0.550	0.774	1.049
1903	0.333	0.565	0.806	1.124
1904	0.323	0.587	0.797	1.040
1905	0.362	0.619	0.871	1.208
1906	0.436	0.650	0.994	1.210
1907	0.397	0.604	0.888	1.106
1908	0.331	0.528	0.757	0.940
1909	0.334	0.514	0.769	1.099
1910	0.383	0.534	0.802	0.962

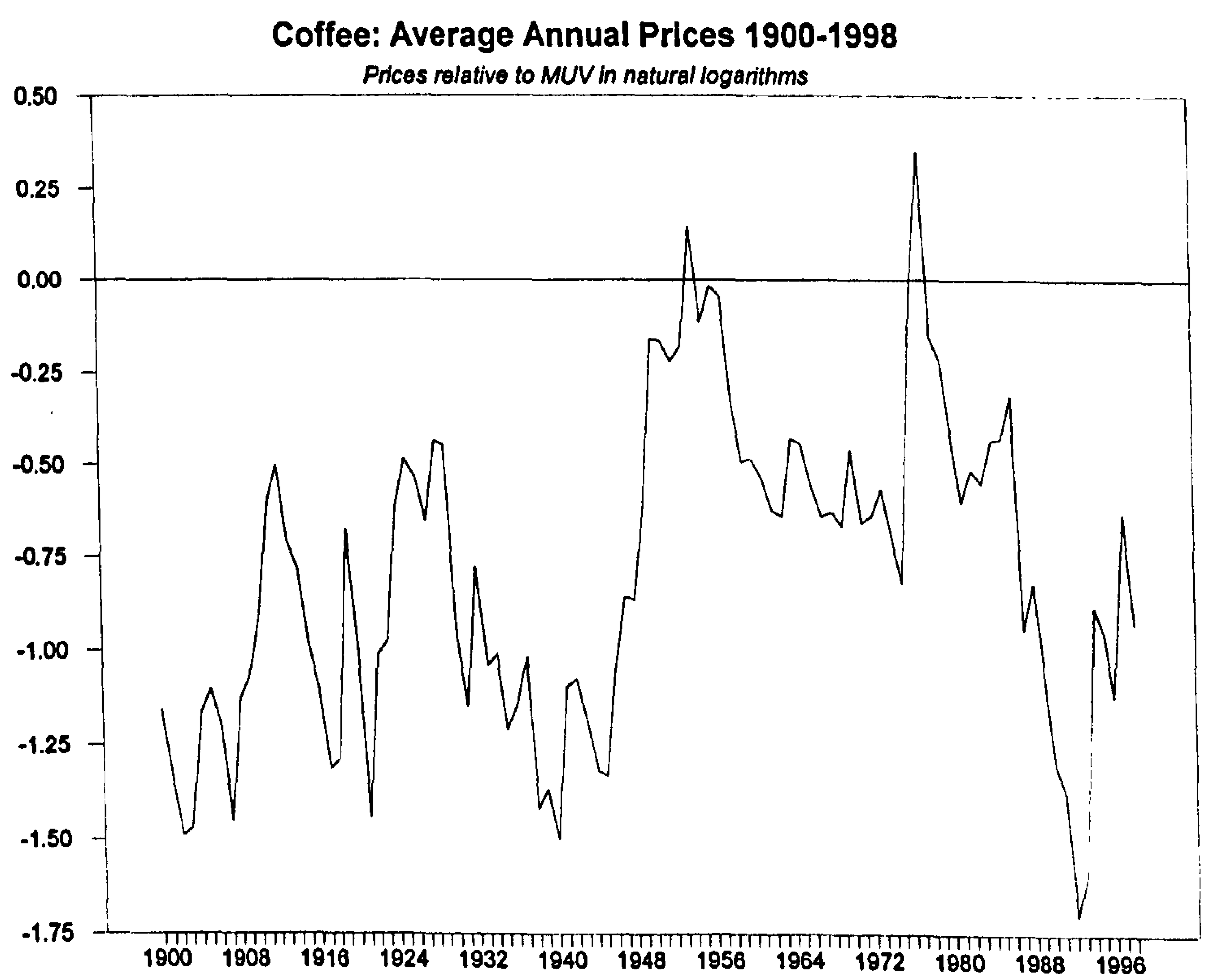
1911	0.475	0.533	0.796	1.082
1912	0.505	0.592	0.785	1.361
1913	0.484	0.582	0.767	1.100
1914	0.396	0.562	0.714	1.067
1915	0.434	0.496	0.841	2.680
1916	0.395	0.530	0.999	2.097
1917	0.471	0.552	1.075	1.228
1918	0.557	0.540	0.746	0.905
1919	0.375	0.586	0.548	0.758
1920	0.268	0.497	0.708	0.777
1921	0.196	0.366	0.478	0.559
1922	0.240	0.442	0.676	0.770
1923	0.314	0.424	0.858	0.889
1924	0.369	0.437	0.956	0.853
1925	0.419	0.444	1.047	1.008
1926	0.498	0.421	1.029	1.023
1927	0.518	0.404	0.873	0.919
1928	0.406	0.467	0.814	0.888
1929	0.378	0.394	0.917	0.996
1930	0.271	0.289	0.756	0.712
1931	0.255	0.266	0.708	0.693
1932	0.276	0.311	0.640	0.661
1933	0.439	0.347	0.697	0.828
1934	0.495	0.404	0.589	0.721
1935	0.489	0.554	0.632	0.768
1936	0.451	0.389	0.733	0.869
1937	0.515	0.378	0.914	1.131
1938	0.384	0.349	0.690	0.765
1939	0.500	0.345	0.804	0.927
1940	0.453	0.281	0.755	1.053
1941	0.444	0.264	0.793	1.166
1942	0.383	0.251	0.765	1.110
1943	0.341	0.261	0.685	0.990
1944	0.300	0.229	0.601	0.870
1945	0.292	0.259	0.586	0.847
1946	0.302	0.395	0.721	0.885
1947	0.358	0.293	1.080	0.881
1948	0.446	0.297	1.300	1.116
1949	0.476	0.307	1.182	1.064
1950	0.504	0.347	1.124	1.336
1951	0.565	0.353	1.248	1.463
1952	0.525	0.329	1.151	1.292
1953	0.435	0.344	0.983	0.902
1954	0.426	0.352	1.046	0.906

1955	0.435	0.363	1.115	1.032
1956	0.446	0.355	1.130	1.085
1957	0.419	0.352	1.024	0.908
1958	0.419	0.348	0.855	0.829
1959	0.449	0.357	0.862	0.921
1960	0.472	0.364	0.669	1.077
1961	0.517	0.361	0.585	0.918
1962	0.512	0.416	0.502	0.778
1963	0.529	0.500	0.578	0.909
1964	0.708	0.497	0.907	1.365
1965	0.801	0.493	1.030	1.301
1966	0.711	0.476	0.820	1.139
1967	0.650	0.565	0.709	1.090
1968	0.621	0.789	0.750	1.057
1969	0.646	0.625	0.857	1.098
1970	0.651	0.582	0.848	1.062
1971	0.590	0.482	0.672	1.055
1972	0.577	0.482	0.733	1.182
1973	0.64	0.631	0.901	2.302
1974	0.895	0.953	1.020	2.751
1975	0.676	0.805	0.645	1.484
1976	0.736	0.782	0.679	1.403
1977	0.954	0.756	0.857	1.060
1978	0.988	0.768	0.798	0.924
1979	1.046	1.392	1.288	1.021
1980	1.036	2.361	0.881	0.955
1981	0.871	1.199	0.704	1.057
1982	0.801	0.920	0.537	0.945
1983	0.830	1.355	0.428	0.992
1984	0.799	0.985	0.455	1.223
1985	0.748	0.737	0.399	1.031
1986	0.339	0.557	0.351	0.842
1987	0.333	0.650	0.470	0.812
1988	0.329	0.565	0.482	1.176
1989	0.401	0.479	0.497	1.583
1990	0.270	0.397	0.567	1.366
1991	0.243	0.325	0.382	0.987
1992	0.254	0.304	0.355	1.050
1993	0.216	0.333	0.267	0.817
1994	0.220	0.395	0.348	0.817
1995	0.231	0.358	0.370	0.780
1996	0.240	0.374	0.475	0.812
1997	0.232	0.373	0.404	1.100
1998	0.237	0.438	0.356	0.890

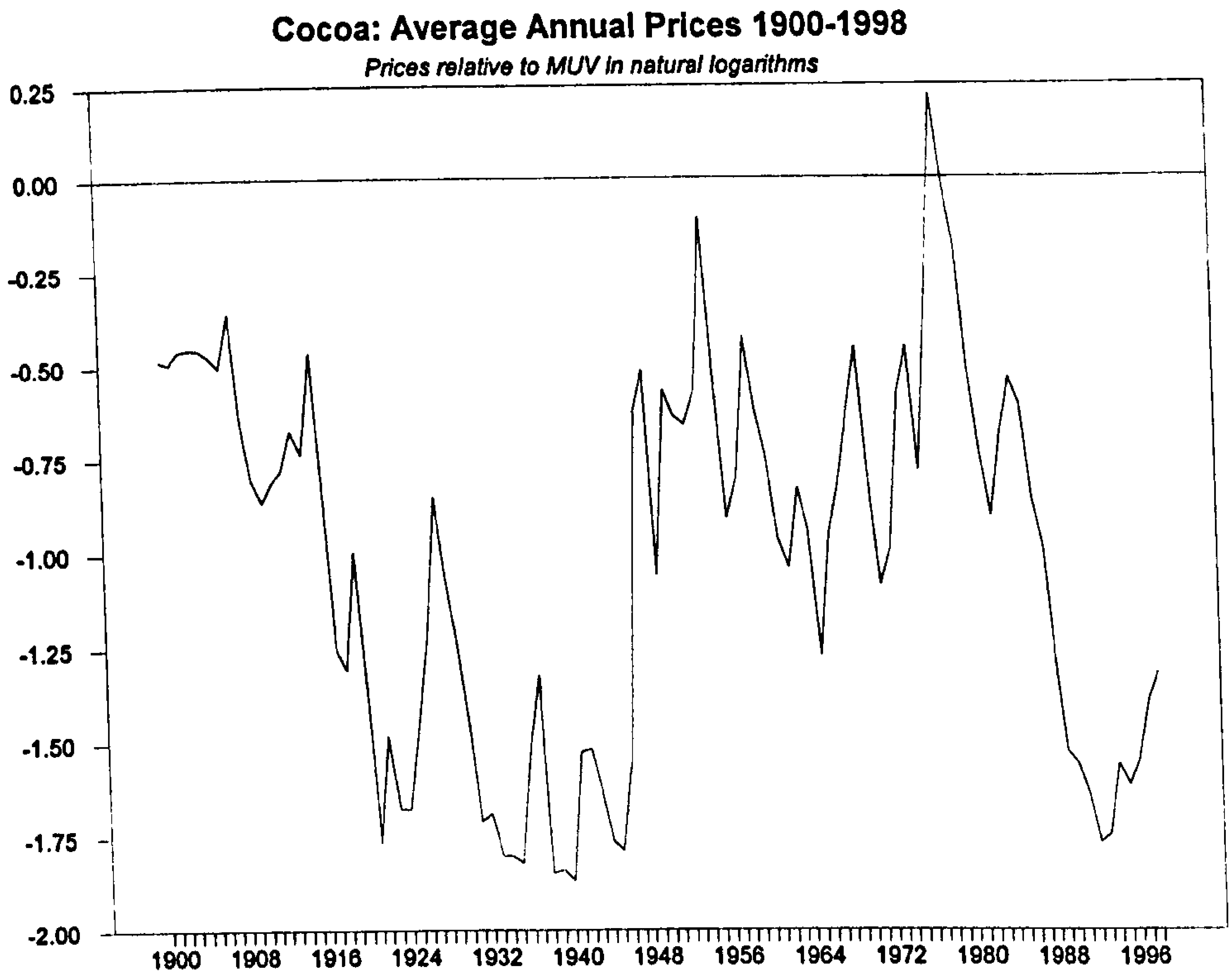
Appendix II.ii. Graphical Illustrations of the Relative Commodity Price Series

The commodity price series described in chapter two are illustrated here. The figures below show the trajectory of relative primary commodity prices in natural logarithms over time.

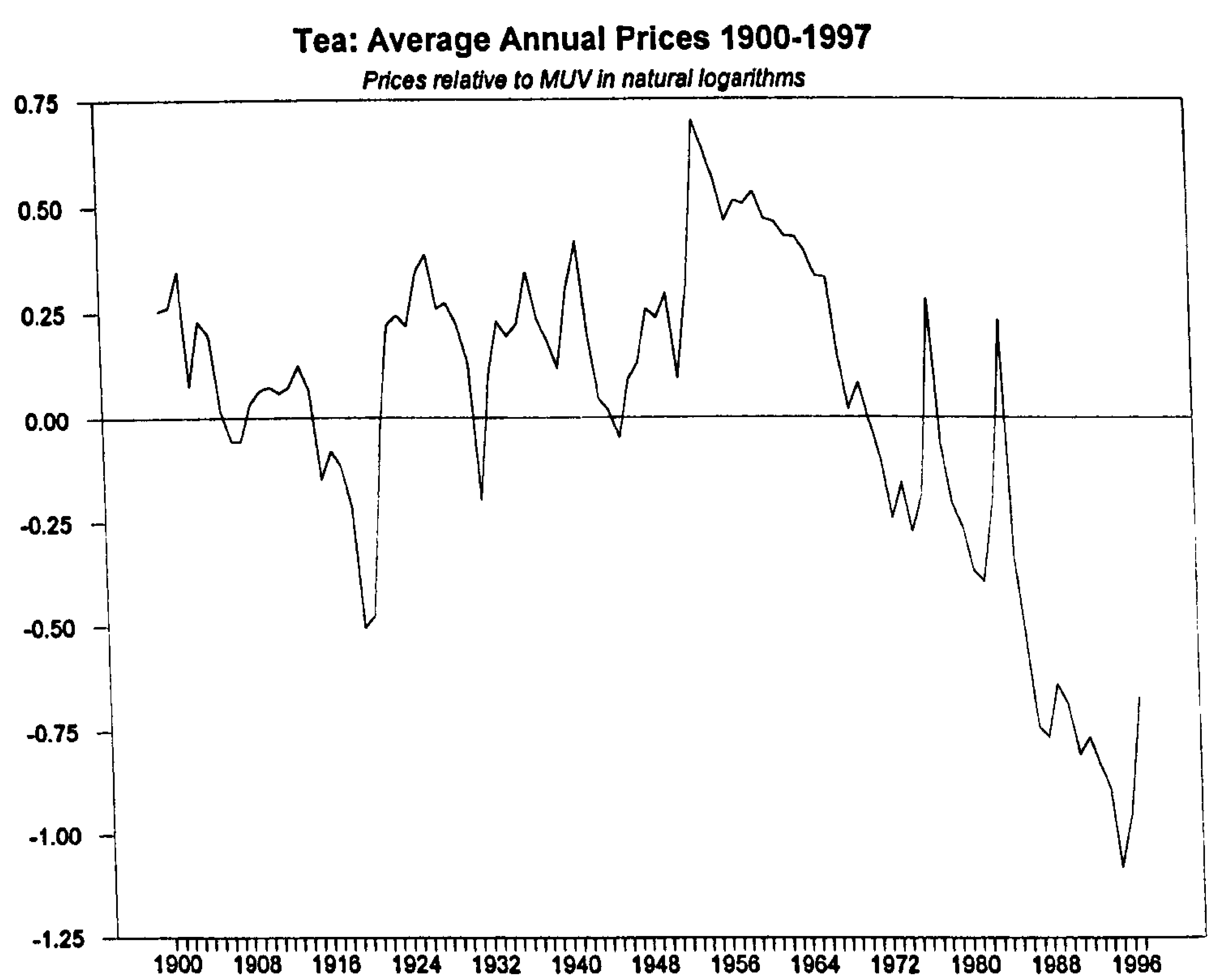
Coffee:



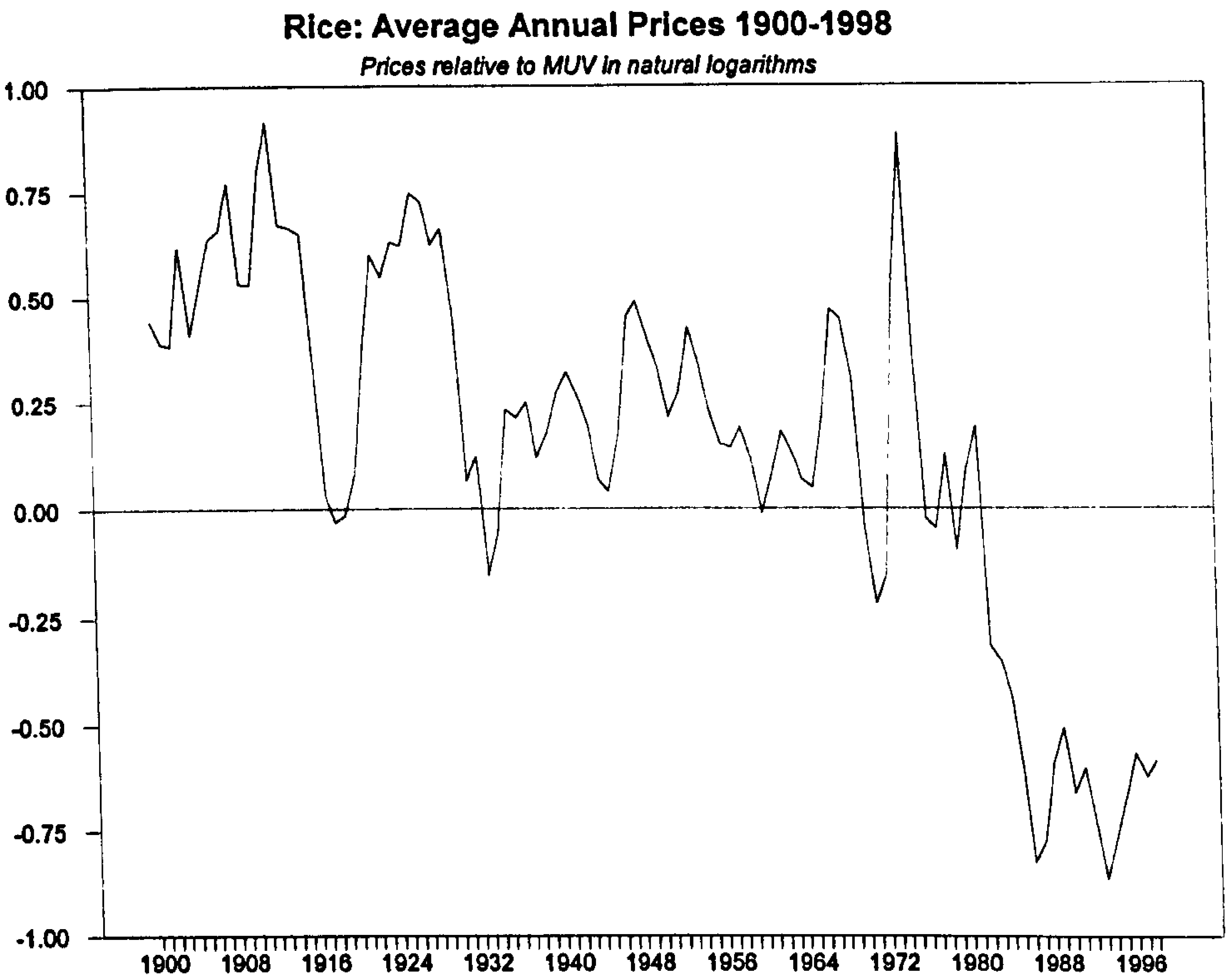
Cocoa:


Cocoa: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
0.25
0.00
-0.25
-0.50
-0.75
-1.00
-1.25
-1.50
-1.75
-2.00
1900 1908 1916 1924 1932 1940 1948 1956 1964 1972 1980 1988 1996

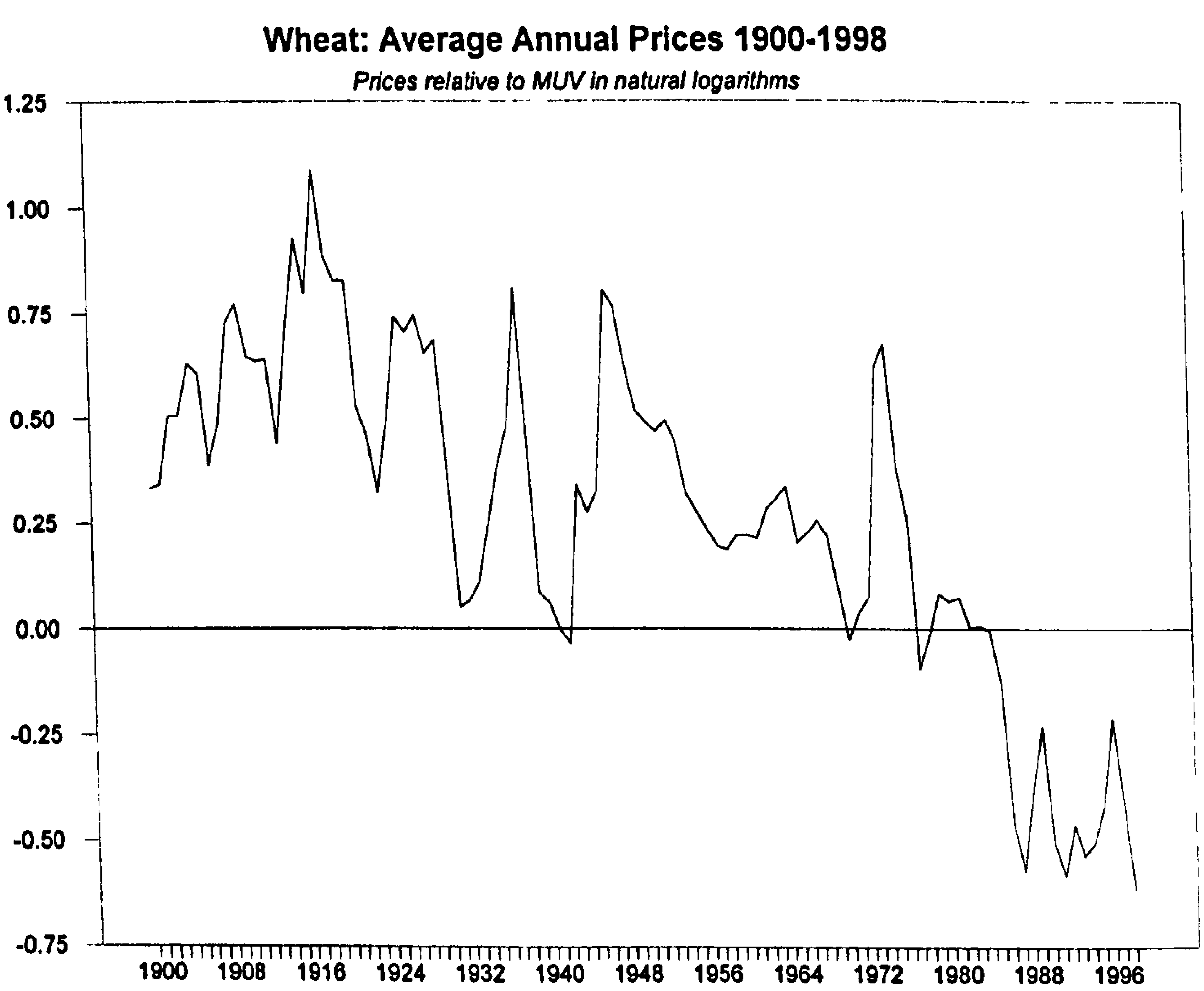
Tea:


Tea: Average Annual Prices 1900-1997
Prices relative to MUV in natural logarithms
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
-1.25
1900 1908 1916 1924 1932 1940 1948 1956 1964 1972 1980 1988 1996

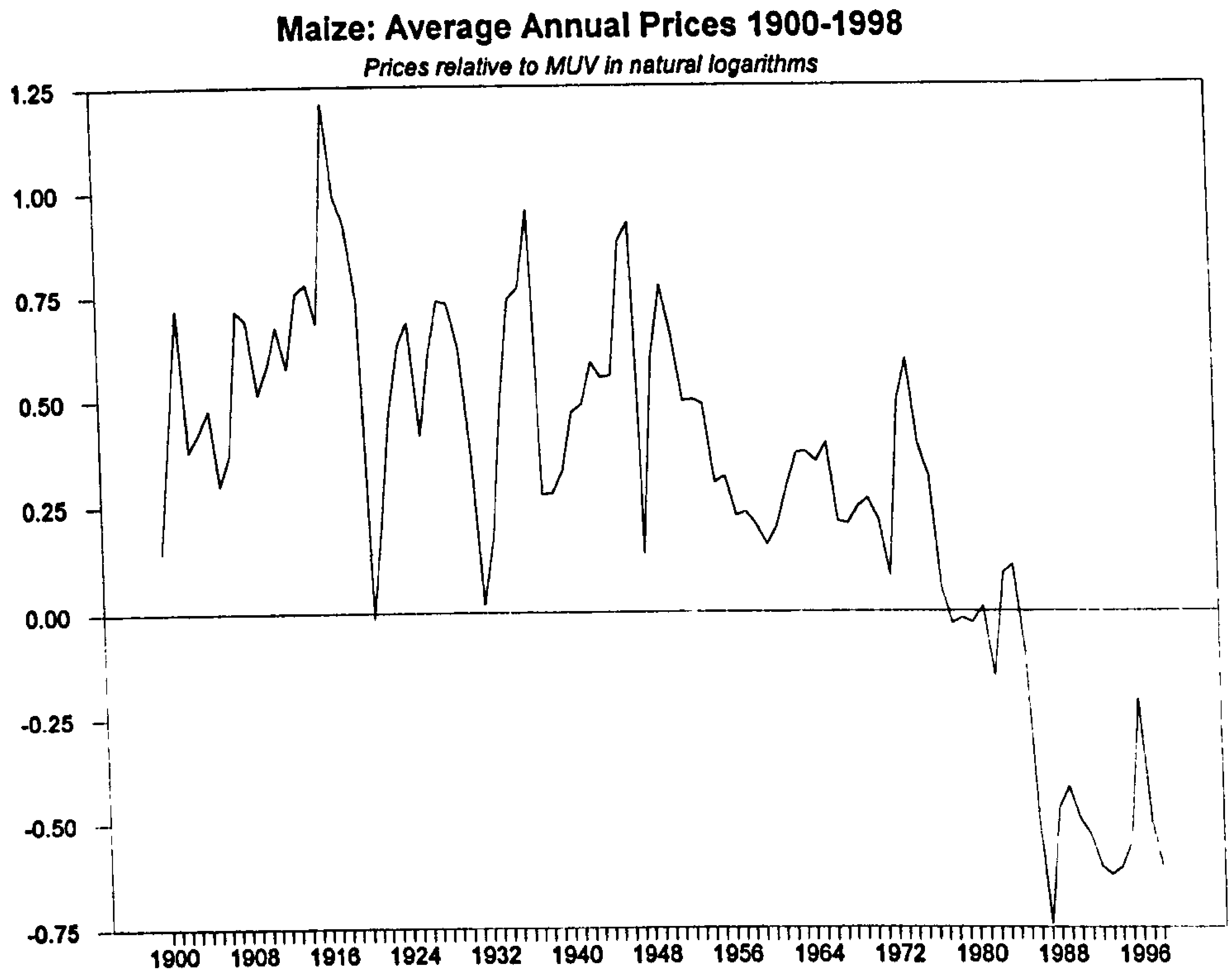
Rice:


Rice: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
1900 1908 1916 1924 1932 1940 1948 1956 1964 1972 1980 1988 1996

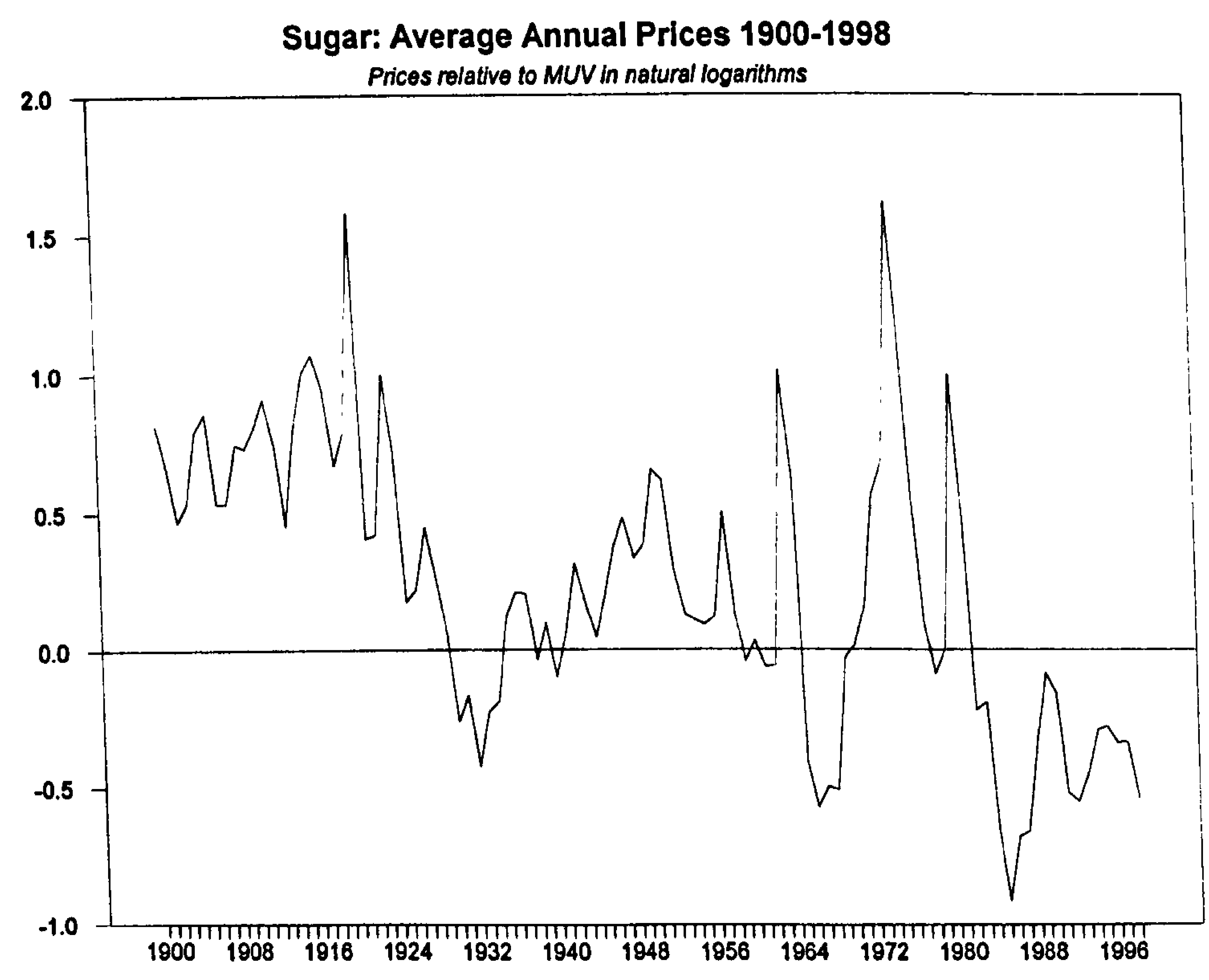
Wheat:


Wheat: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
1.25
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
1900 1908 1916 1924 1932 1940 1948 1956 1964 1972 1980 1988 1996

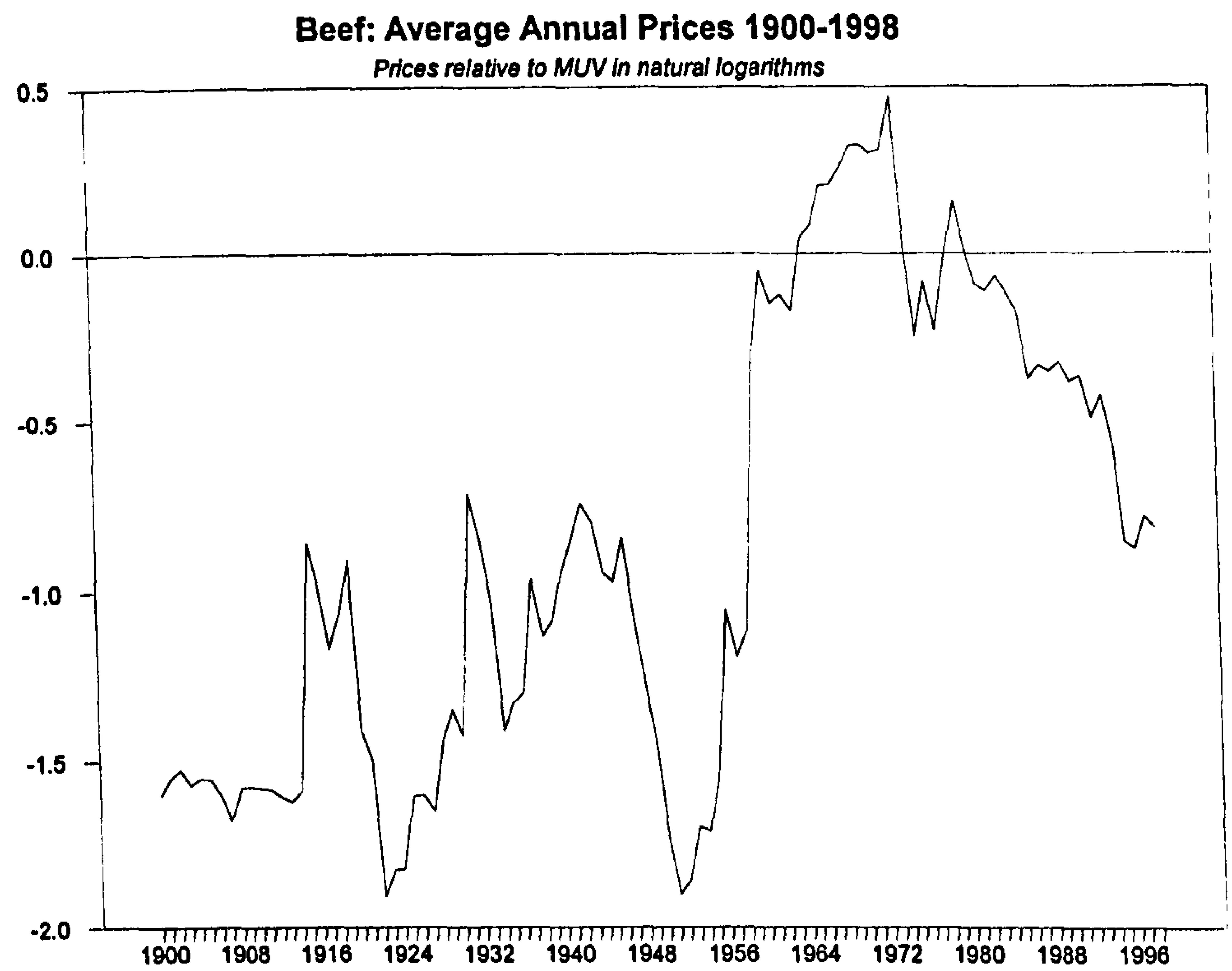
Maize:


Maize: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
1.25
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
1900
1908
1916
1924
1932
1940
1948
1956
1964
1972
1980
1988
1996

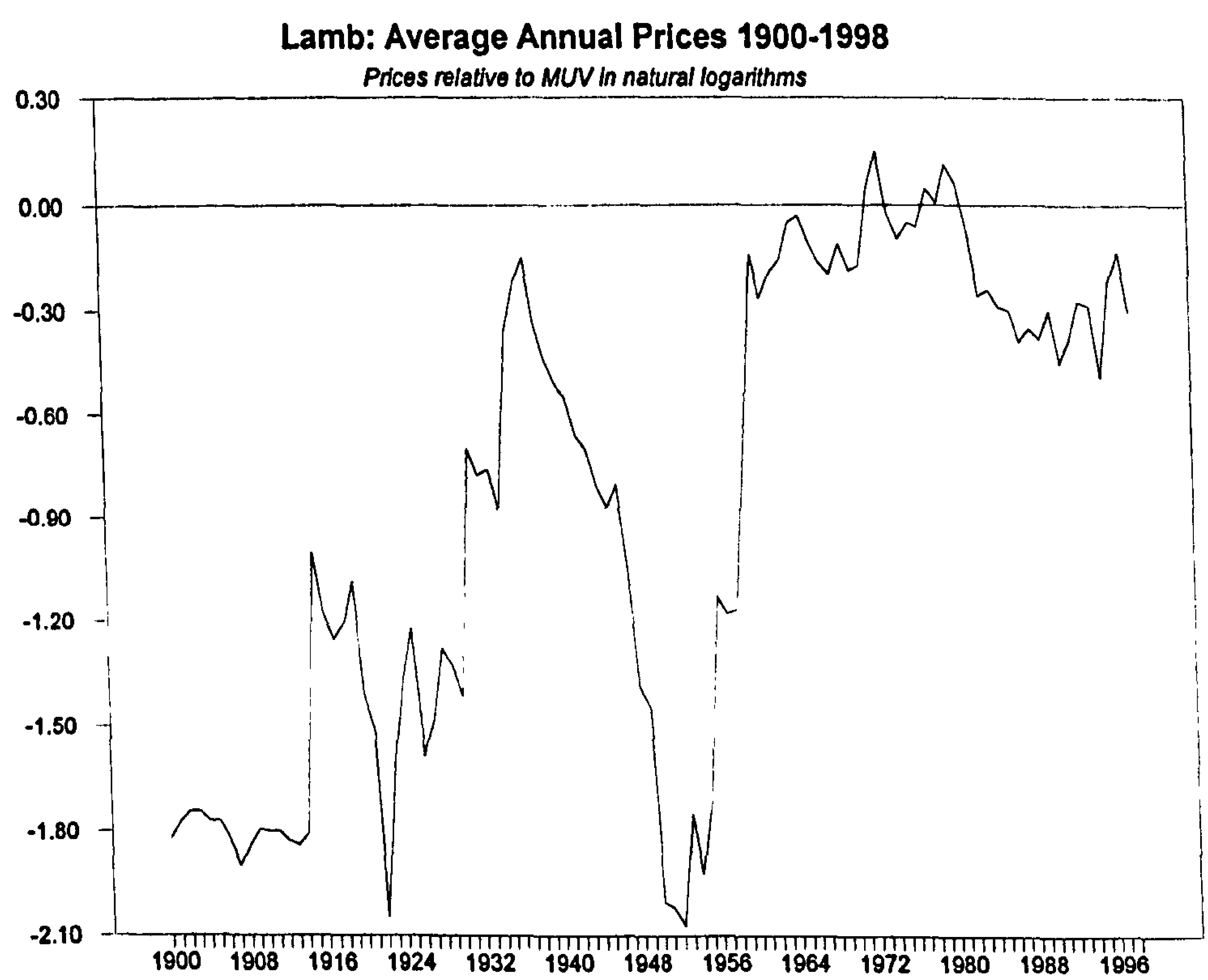
Sugar:


Sugar: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
1900
1908
1916
1924
1932
1940
1948
1956
1964
1972
1980
1988
1996

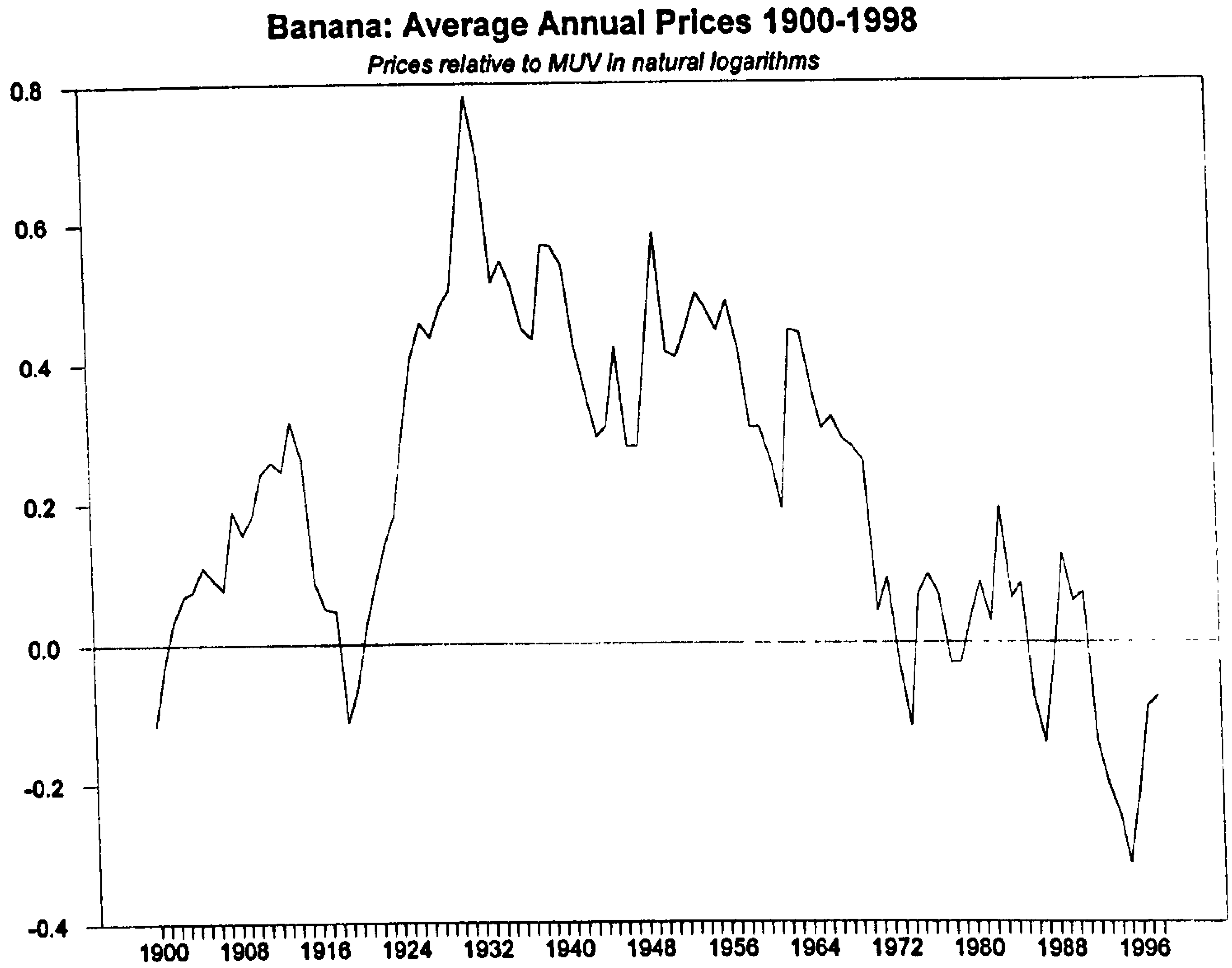
Beef:


Beef: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
0.5
0.0
-0.5
-1.0
-1.5
-2.0
1900
1908
1916
1924
1932
1940
1948
1956
1964
1972
1980
1988
1996

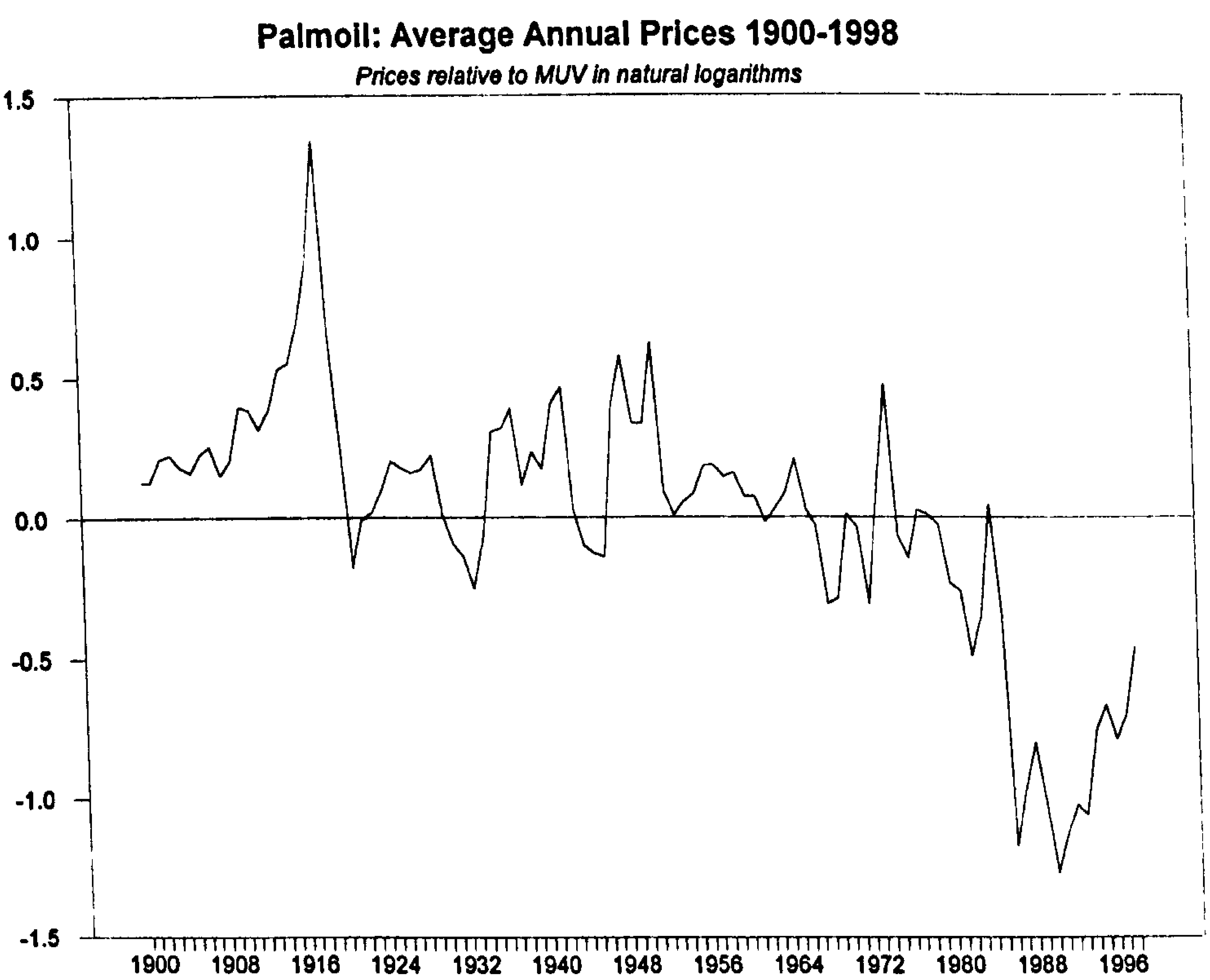
Lamb:


Lamb: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
0.30
0.00
-0.30
-0.60
-0.90
-1.20
-1.50
-1.80
-2.10
1900
1908
1916
1924
1932
1940
1948
1956
1964
1972
1980
1988
1996

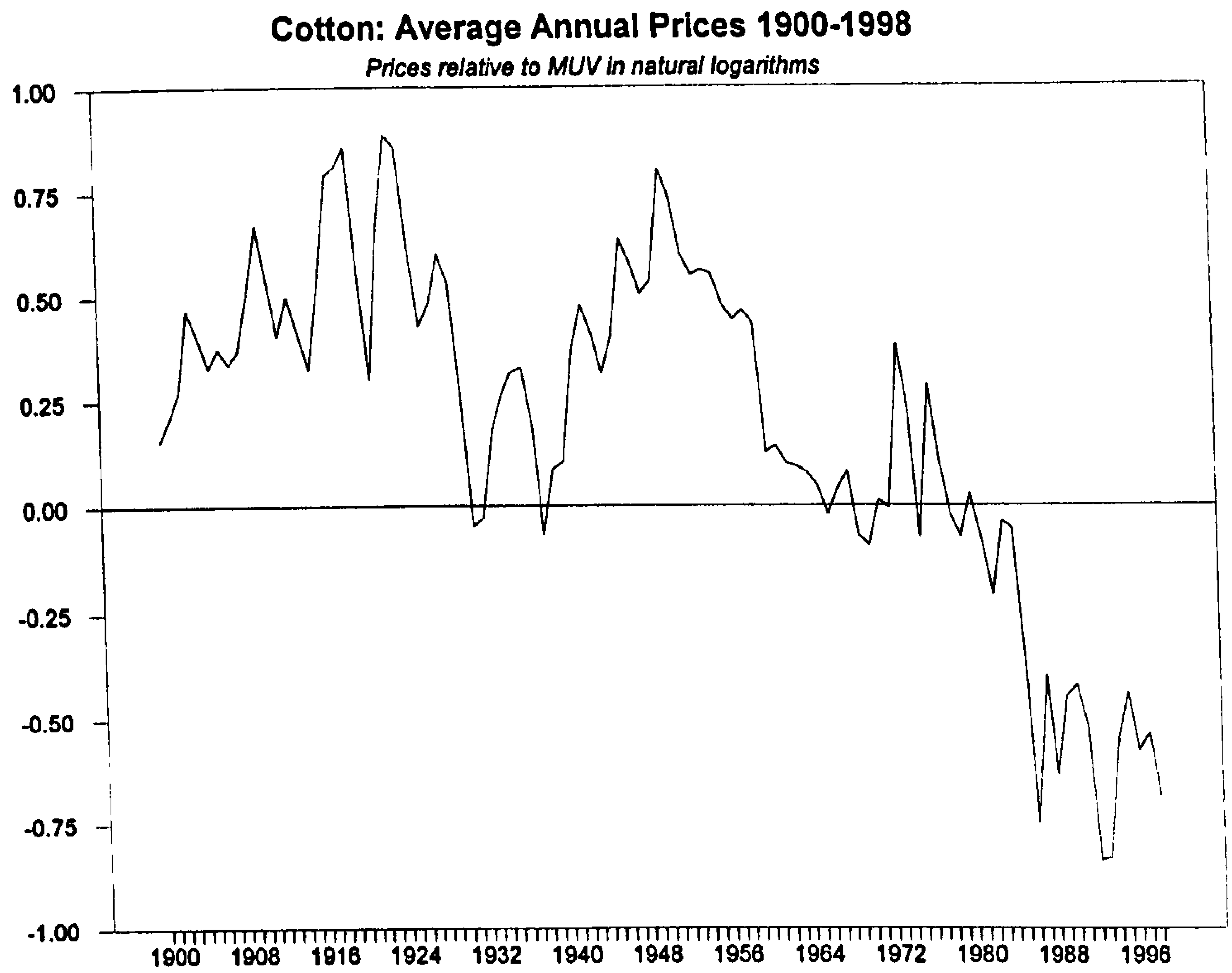
Banana:



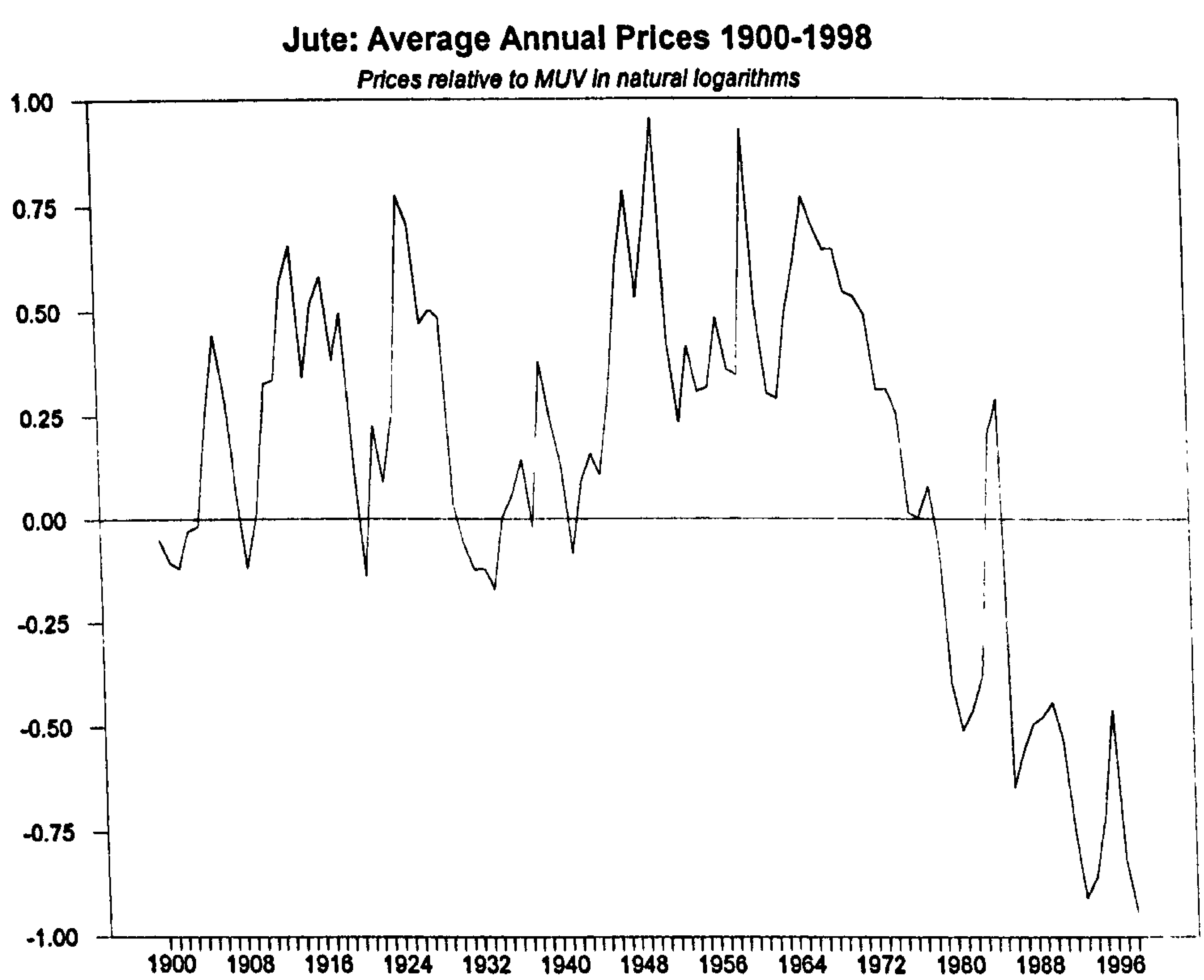
Palm Oil:



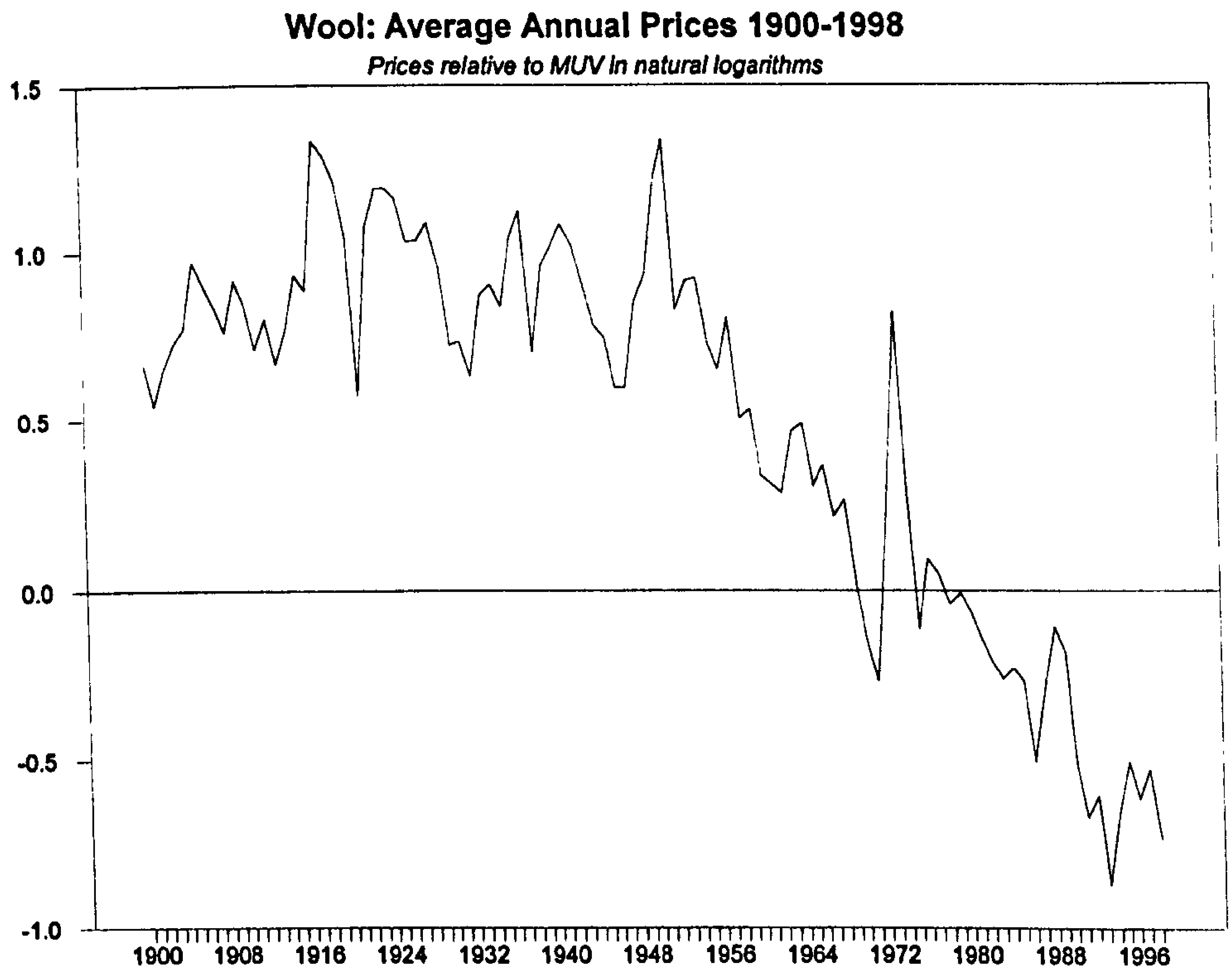
Cotton:


Cotton: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
1900 1908 1916 1924 1932 1940 1948 1956 1964 1972 1980 1988 1996

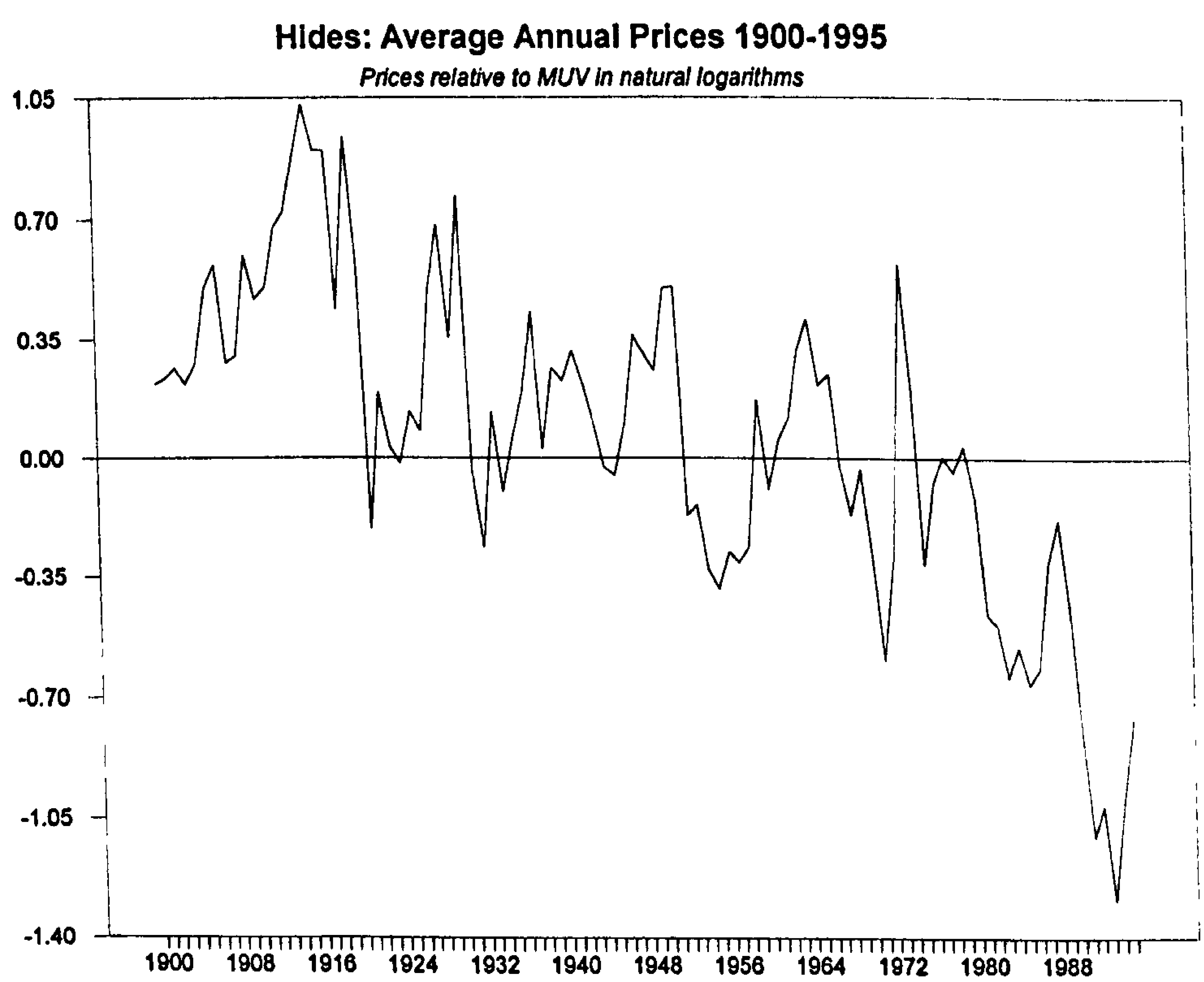
Jute:


Jute: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
1900 1908 1916 1924 1932 1940 1948 1956 1964 1972 1980 1988 1996

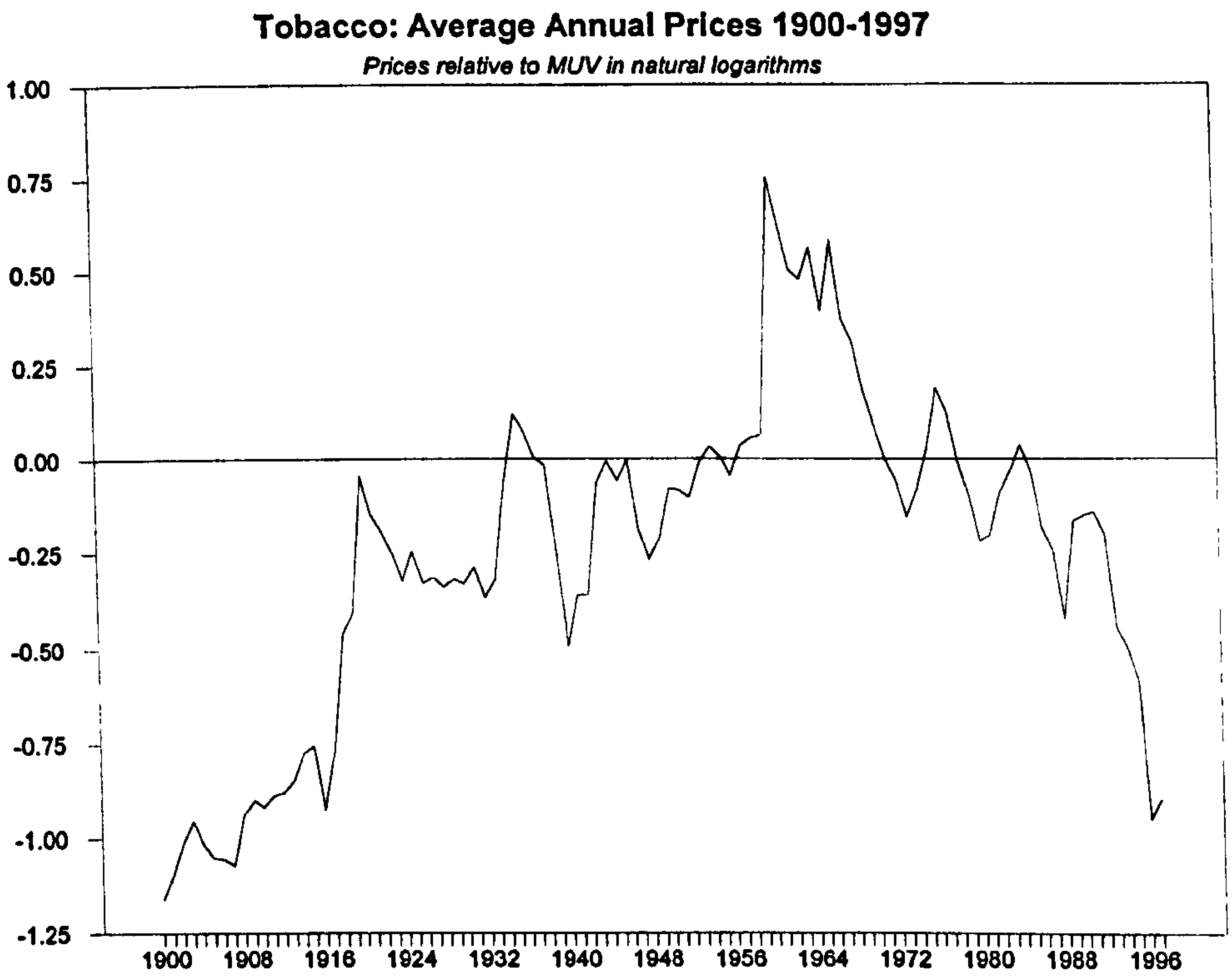
Wool:


Wool: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
1.5
1.0
0.5
0.0
-0.5
-1.0
1900
1908
1916
1924
1932
1940
1948
1956
1964
1972
1980
1988
1996

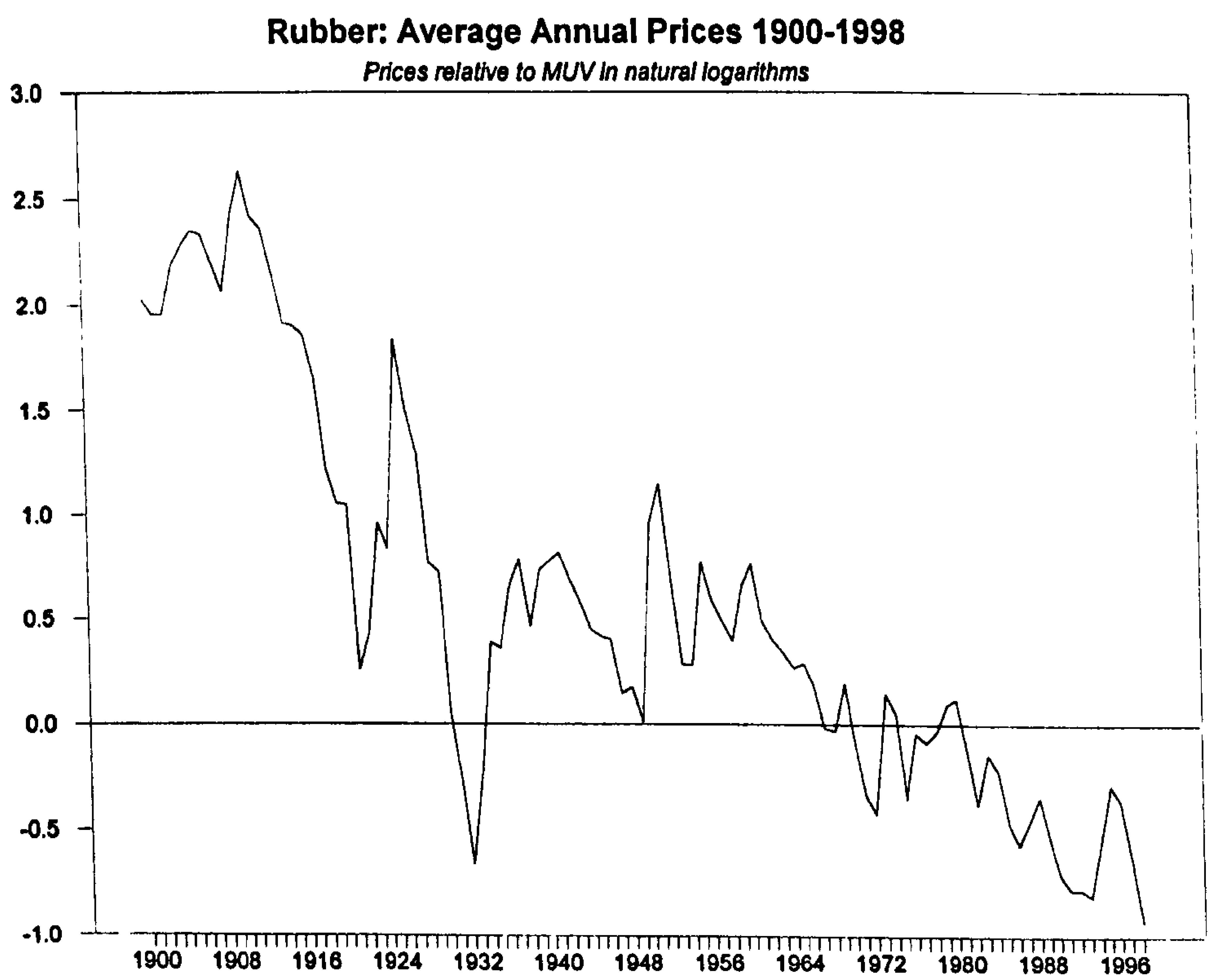
Hides:


Hides: Average Annual Prices 1900-1995
Prices relative to MUV in natural logarithms
1.05
0.70
0.35
0.00
-0.35
-0.70
-1.05
-1.40
1900
1908
1916
1924
1932
1940
1948
1956
1964
1972
1980
1988

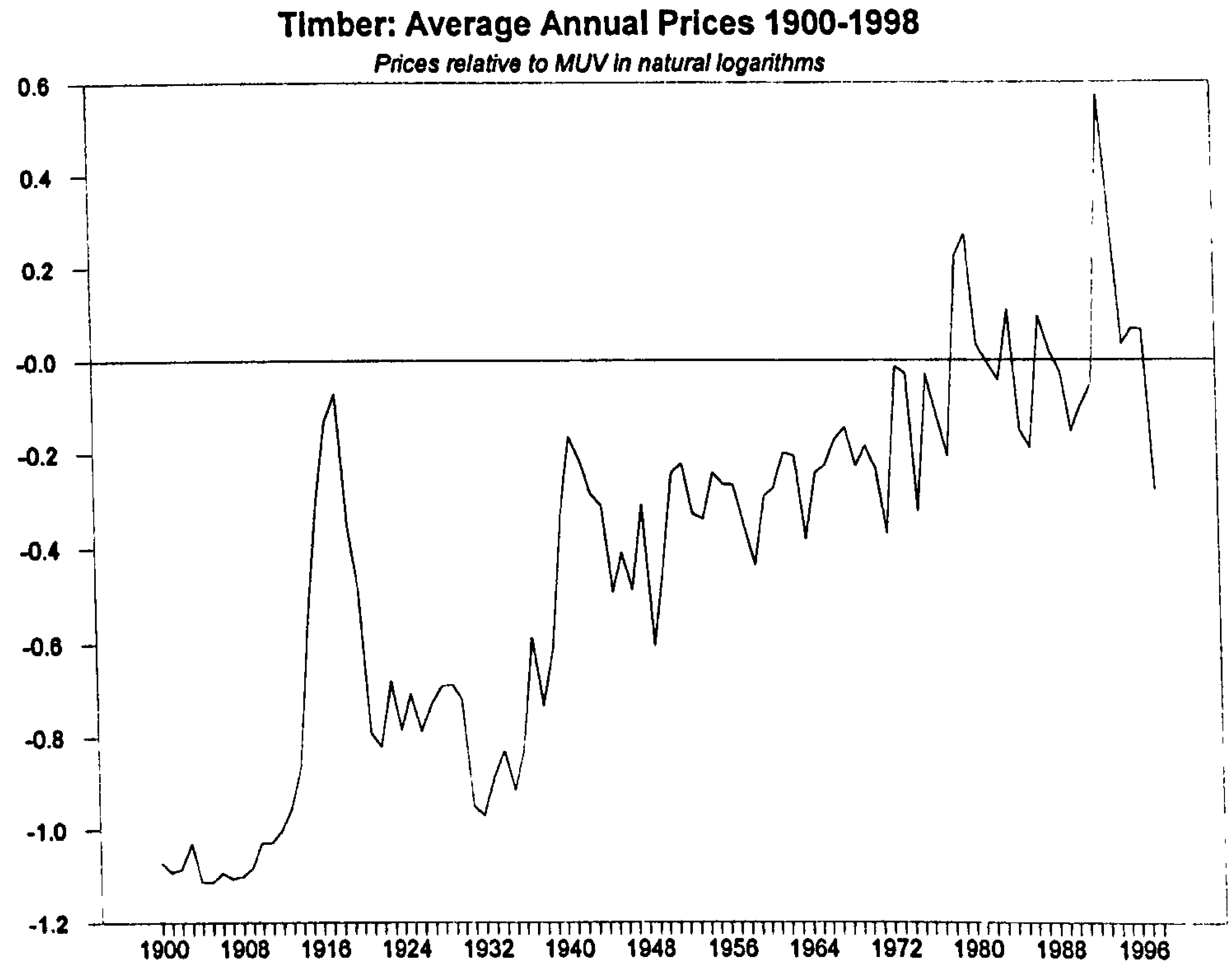
Tobacco:



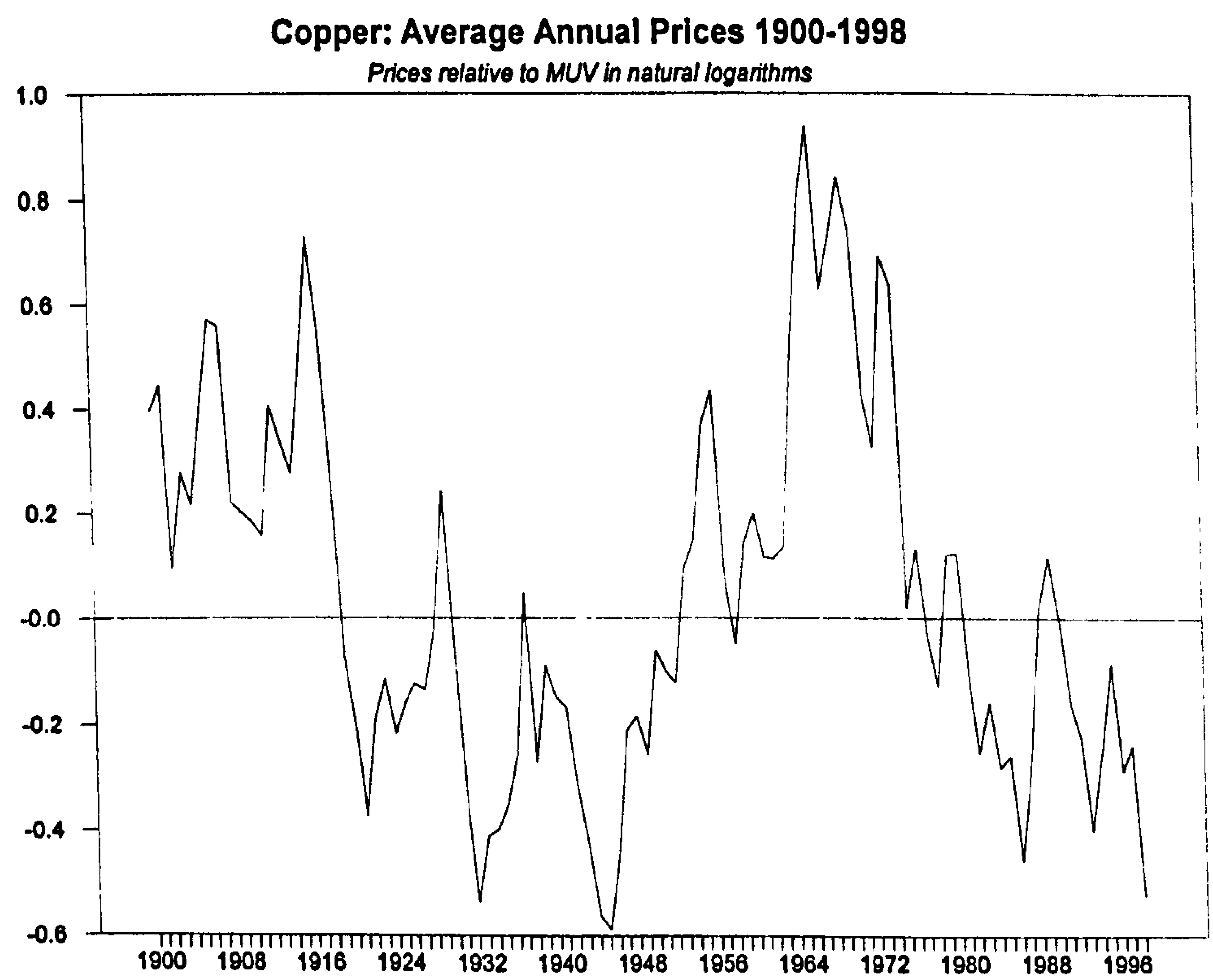
Rubber:



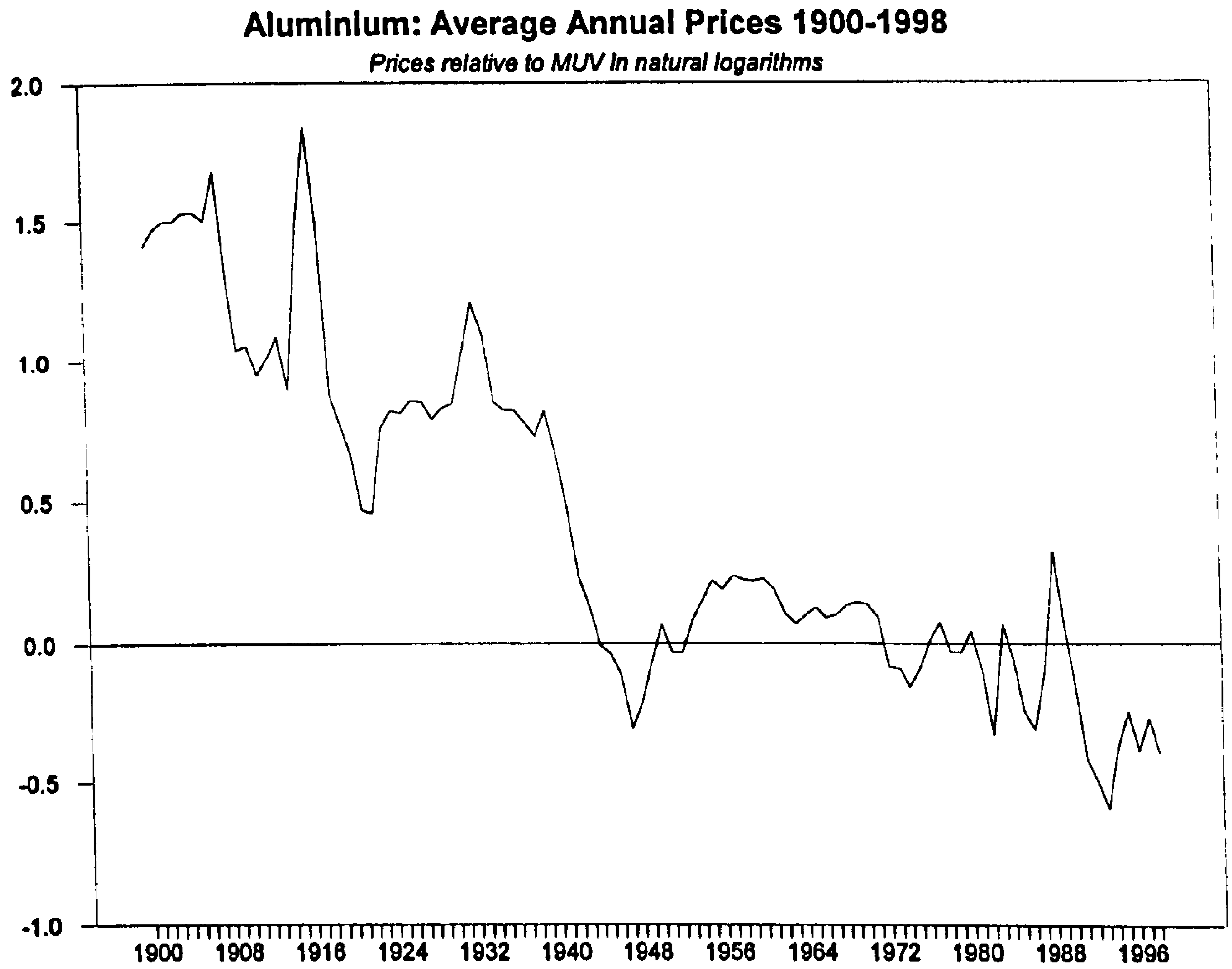
Timber:



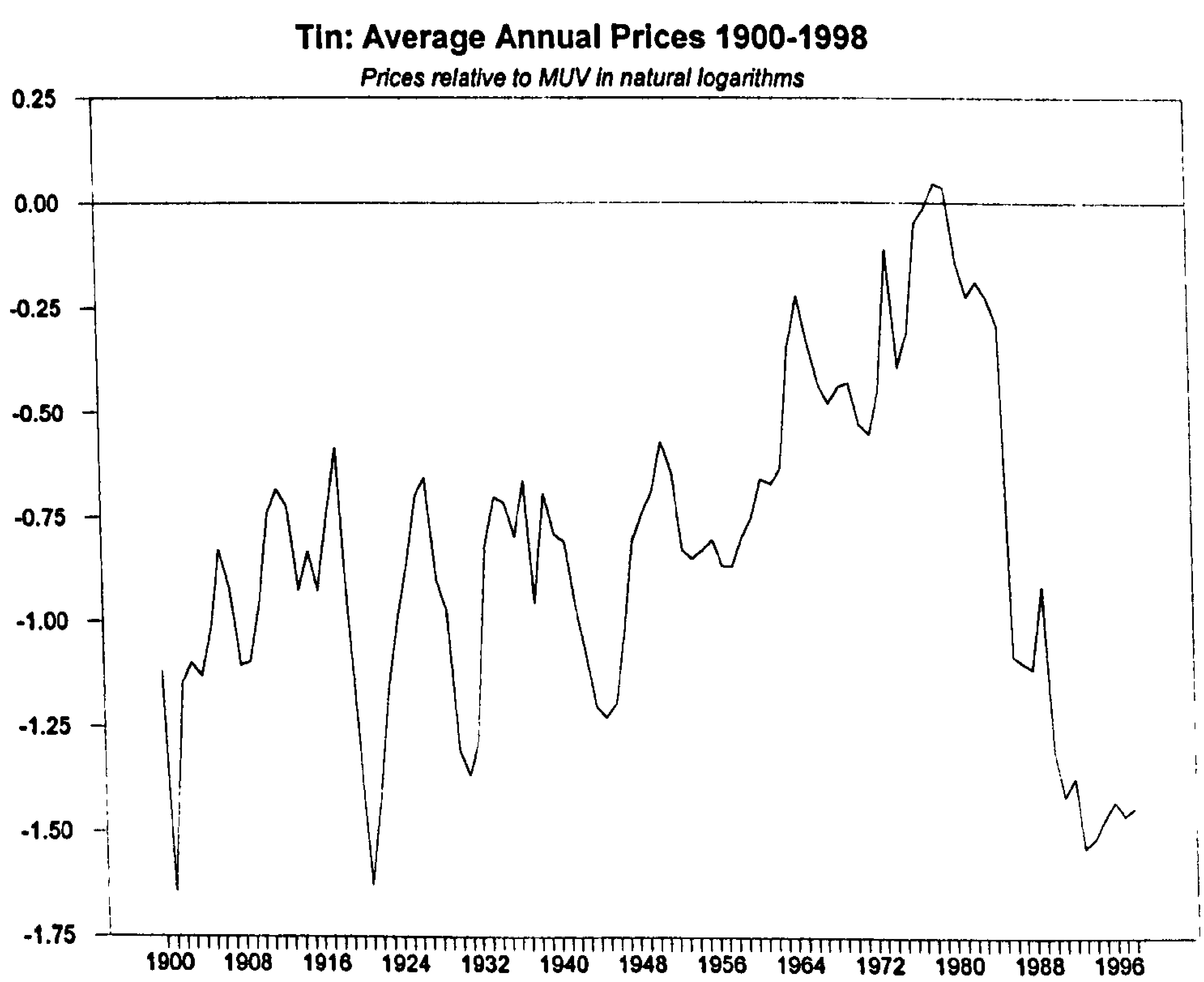
Copper:



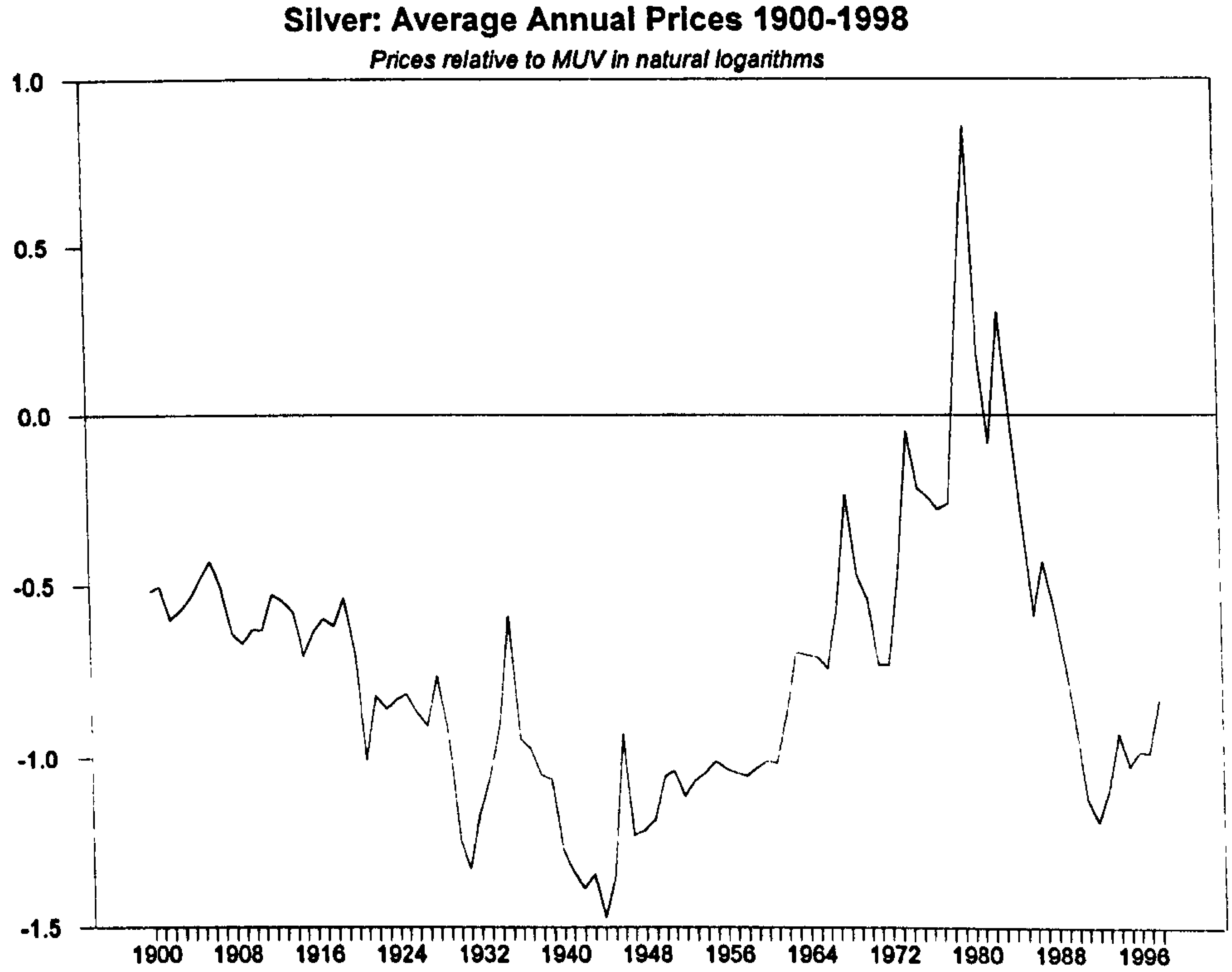
Aluminium:



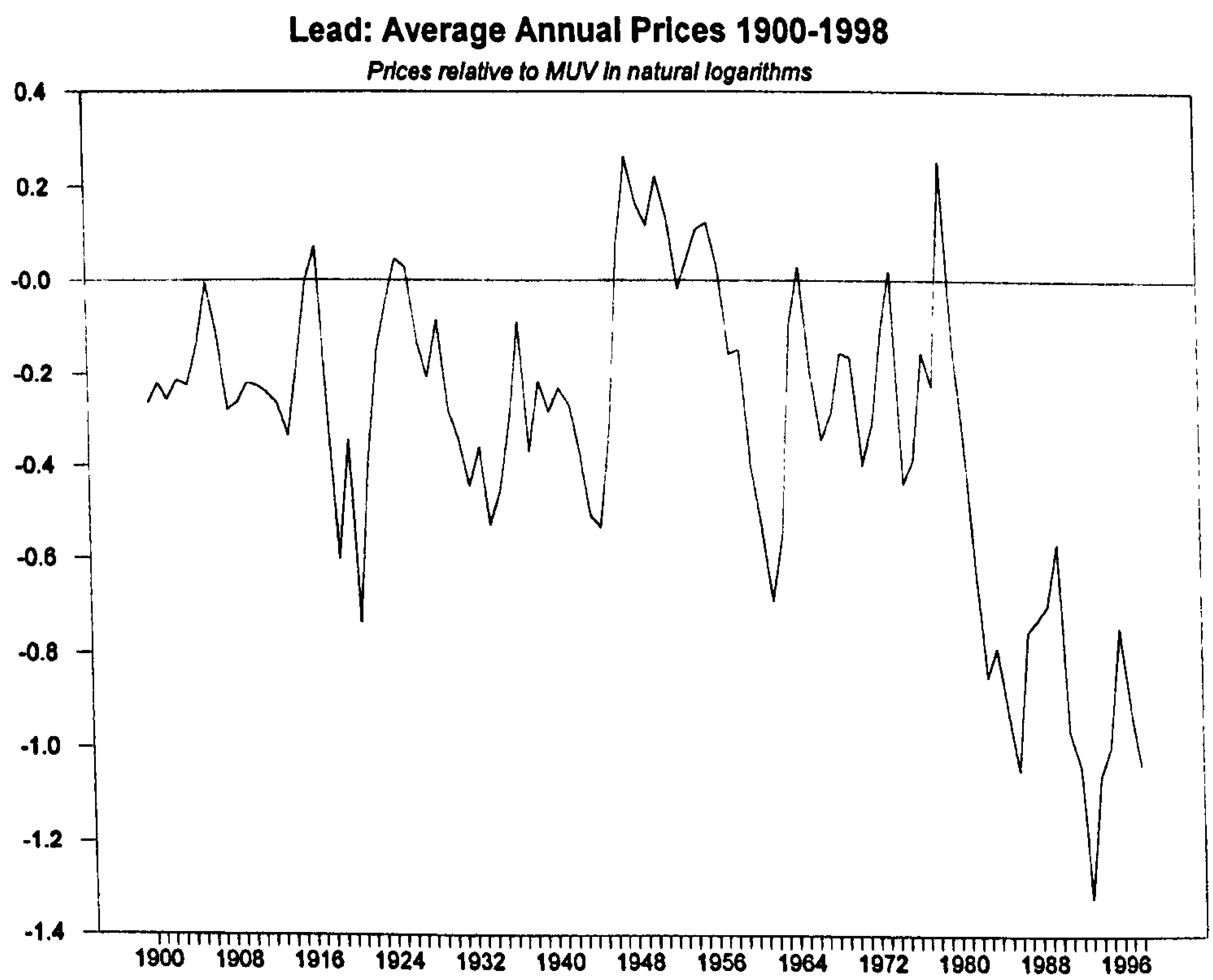
Tin:



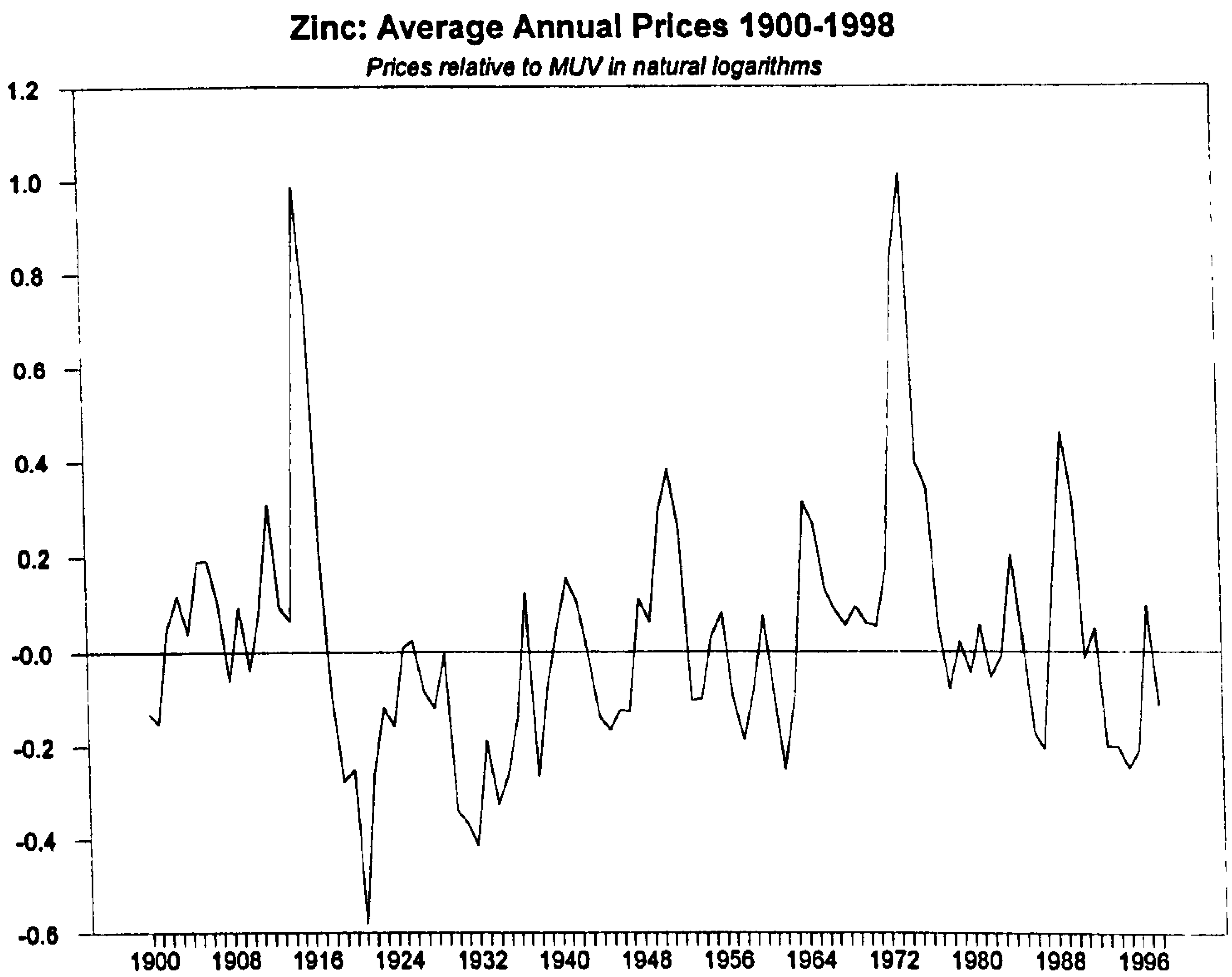
Silver:


Silver: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
1.0
0.5
0.0
-0.5
-1.0
-1.5
1900 1908 1916 1924 1932 1940 1948 1956 1964 1972 1980 1988 1996

Lead:


Lead: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
0.4
0.2
-0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
-1.4
1900 1908 1916 1924 1932 1940 1948 1956 1964 1972 1980 1988 1996

Zinc:


Zinc: Average Annual Prices 1900-1998
Prices relative to MUV in natural logarithms
1.2
1.0
0.8
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
1900
1908
1916
1924
1932
1940
1948
1956
1964
1972
1980
1988
1996

Chapter 3

Estimation Results for the Trend Coefficient

Chapter 3: Estimation Results for the Trend Coefficient

3.1. Inference on stationarity

It is standard practice in applied econometric research to test time series data for stationarity using conventional unit root tests such as the Dickey-Fuller test or the Augmented Dickey Fuller (ADF) test (*cf.* Enders (1995)). The ADF test with an appropriate number of autoregressive lags was applied to all commodities other than Hides[1], using general to specific testing as recommended in Enders (1995). The initial testing equation included a total of five lags for all commodities. Insignificant lags were then eliminated one at a time, and the equation was re-estimated with the reduced number of lags. This process was repeated until the last lag in the testing equation was significant or the testing equation reduced to the ordinary Dickey Fuller test allowing for trend and constant.

As a result of this testing procedure, only five commodities (Sugar, Lamb, Timber, Aluminium and Zinc) are classified as trend stationary, while the remainder (Coffee, Cocoa, Tea, Rice, Wheat, Maize, Beef, Bananas, Palm Oil, Cotton, Jute, Wool, Rubber, Tobacco, Copper, Tin, Silver and Lead) are identified as stationary in first differences. (Details of the test results for the commodities mentioned above are given in appendix III.i).

It is pointed out in Enders (*op. cit.*) that the above mentioned testing procedure will identify the correct lag length if the lag length for the most general initial Dickey Fuller testing equation contains at least the correct number of lagged terms or

[1] This series has been dropped since data are not available after 1995.

more, and if the true generating process is purely autoregressive. However, neither of these two conditions should be taken for granted. Even if the generating process is purely autoregressive, the number of lags one can in practice include in the equation is often limited by the size of the data set available. In this context it has been shown by Agiakloglou and Newbold (1992) that a large number of autoregressive lags can lead to an unnecessary loss in power. Another serious problem in practice is the possibility of large moving average components in the generating process. Agiakloglou and Newbold (1992) show that a moving average parameter close to the invertibility boundary, can lead to spurious rejections of the null hypothesis. If a deterministic trend term is included in the testing equation, this problem can occur even for relatively low absolute values of moving average coefficients (see Agiakloglou and Newbold (1996)).

As discussed in Chapter 2 and in Newbold and Vougas (1996), there are reasons to be sceptical about the quality of inferences on stationarity solely on the basis of unit root tests. For this reason, no *ex ante* model selection is made here on the basis of ADF or other unit root tests. Rather, trend and difference stationary models will be estimated for all the commodities under consideration, a method which should also allow for an assessment of the impact of model specification on conclusions concerning the significance and magnitude of trend estimates.

3.2 Estimation Results for ARIMA models

For all the commodities for which adequate data coverage was available[2], univariate series for all possible configurations with ARMA(p,0,q) such that $p+q \leq 5$, were estimated in levels using exact maximum likelihood estimation. In addition to a linear trend term, a constant was included in all models. Among these estimates, models were selected assuming the lowest value for the Schwarz-Bayesian Criterion (SBC) to indicate the most adequate model configuration. For the difference stationary case, ARIMA(p,1,q) models with $p+q \leq 5$ were estimated, again using exact maximum likelihood methods as above. Inferences about the drift term are in this case made from the estimated coefficient on the constant for each commodity. As in the previous case, the Schwarz Bayesian Criterion was used for model selection. Full estimation results for the selected estimates are reported in appendix III.ii for the trend stationary models and appendix III.iii for the difference stationary models[3].

An interesting result with respect to the debate concerning the Prebisch-Singer Hypothesis as well as with respect to a possible extrapolation of price developments is the evidence on the presence of significant trend terms. The present chapter therefore concentrates on estimates giving evidence in favour of the presence or absence of secular trends as well as on possible indications of the

---

[2] The series for Tea and Tobacco extend over the period 1990-1997, whereas the remaining 21 data series cover the period 1900-1998. The data series for Hides could only be obtained for the period 1900-1995, and has been omitted from most of the present study since there would be little scope for comparison with the remaining data series.

[3] These estimation results provide evidence of overdifferencing in two cases (Aluminium and Zinc). Since no *a priori* conclusions on the order of integration are used for model selection at this stage, the issue of overdifferencing will be taken up later in Chapter 5.

magnitude of such trends. Estimated coefficients and t-ratios for the linear trend and for the drift term in difference stationary models are listed in Table 3.2.1.

Table 3.2.1. Estimates of trend and drift coefficients

Commodity	Trend coefficient	t-ratio	Coefficient on the drift term	t-ratio
Coffee	0.004	1.131	0.002	0.093
Cocoa	-0.003	-0.556	-0.009	-0.457
Tea†	-0.007	-1.879	-0.010	-0.573
Rice	-0.011	-5.223	-0.012	-2.423
Wheat	-0.011	-6.866	-0.010	-1.154
Maize	-0.010	-4.180	-0.010	-1.341
Sugar	-0.011	-4.108	-0.012	-0.844
Beef	0.014	2.295	0.008	0.387
Lamb	0.018	5.102	0.015	0.715
Bananas	-0.001	-0.346	0.000	0.043
Palm Oil	-0.010	-4.332	-0.007	-0.429
Cotton	-0.010	-3.328	-0.008	-0.765
Jute	-0.007	-1.566	-0.008	-0.665
Wool	-0.016	-4.652	-0.014	-1.850
Tobacco†	0.005	0.574	0.003	0.189
Rubber	-0.028	-6.764	-0.030	-1.041
Timber	0.011	7.321	0.008	0.493
Copper	-0.004	-1.097	-0.009	-0.490
Aluminium	-0.019	-8.997	-0.019	-7.931
Tin	0.001	0.202	-0.003	-0.172
Silver	0.000	0.050	-0.003	-0.229
Lead	-0.006	-2.166	-0.008	-0.414
Zinc	0.001	0.369	0.000	0.184

† Data Series from 1900-1997 only.

The results in table 3.2.1. suggest that trend stationary and difference stationary models tend to yield similar values for the estimate of a secular trend. The differences in estimated coefficients on linear trend terms in trend stationary models and drift terms in difference stationary models tend to be moderate and the signs of the estimated coefficients are the same in all but three cases: Bananas, Tin and Silver. The estimated trend coefficient on Banana prices is negative if the relative price series is modelled as trend stationary and positive if it is modelled as difference stationary. For Tin and Silver prices the estimated coefficient for the trend is positive in the trend stationary model and negative in the difference stationary model. However, in all those three cases the estimates from either model are highly insignificant.

It is with respect to significance levels that the estimates produced by trend stationary and difference stationary models differ strongly. With the exception of the t-ratio on the -insignificant- estimate of the trend coefficient for Silver the absolute values of the t-ratios for all estimated trend coefficients are lower in the difference stationary than in the corresponding trend stationary case. The fall in significance levels is often very pronounced, in many cases leading to a change from very high significance levels to levels low enough to suggest that coefficient estimates are entirely insignificant. Among the coefficient estimates from trend stationary models, the estimated coefficient is shown to be significant in 13 out of 23 cases at the 5% significance level. The coefficient estimates from the difference stationary model show coefficients to be significant at the 5% level in only two cases (Rice and Aluminium) and even in those cases the significance level has fallen noticeably. The estimated trend coefficient on Wool remains significant at the 10% level.

This discrepancy highlights the importance of the issue of using unit root tests to decide on model specification *a priori*. Using this methodology, it would have appeared that all but one[4] of the price series identified as trend stationary had significant trends, when estimating the models for the relevant price series in levels. On the other hand, no significant trend term would have been found in any of the price series identified as difference stationary, with the exception of Rice, where the models for all these price series would have been estimated in first differences only. (Another possible exception is Wool, where the drift coefficient would have remained significant at a 10 percent level.) Comparing both possible specifications for the price series identified as trend stationary through *a priori* testing, the evidence in favour of significant trend estimates is weakened considerably: only one of these series (Aluminium) retains clear evidence in favour of a significant trend term under the difference stationary specification, while four of the five series presumed trend stationary now have no significant drift terms. On the other hand, nine[5] of the 18 price series identified as difference stationary would have significant trend parameters when modelled as trend stationary[6].

3.2.1. Evaluating trend estimates by commodity groups

Bearing in mind the uncertainties surrounding the validity of unit root tests, as well as the strong implications of test results for conclusions on the significance of trend estimates, it should nevertheless be possible to proceed towards a tentative

[4] The one exception is Zinc, while Sugar, Lamb Aluminium and Timber are shown to have significant trend coefficients.

[5] These are Rice, Wheat, Maize, Beef, Cotton, Wool, Palm Oil, Rubber and Lead.

[6] Higher order ARIMA models may be considered for commodities with large supply response lags (*e.g.* Coffee, Cocoa, Palm Oil or Timber). Re-selecting models by SBC subject to $p+q\leq 9$ only yields a different model parameterisation for Cotton (ARMA(4,2)). For the four commodities mentioned previously higher parametrisations could be justified on the basis of their ACF and PACF plots. This would affect inference on the trend coefficient only for the stationary model of Timber, where the drift coefficient estimate is now significant for ARIMA(0,1,14).

interpretation of the results obtained so far. The price series of the three cereals covered in the data set (Rice, Wheat and Maize) are identified as difference stationary in the unit root tests. They show significant negative trends in the trend stationary model but the evidence on the drift term for the difference stationary alternative is less clear. Table 3.2.2. summarises the results for these three commodities, giving trend estimates as well as standard errors and 95% confidence intervals[7].

Table 3.2.2. Trend Estimates for Cereal Prices

Commodity	Trend Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval	Drift Coeff. (*100)	Stnd Error (*100)	95% Conf. Interval
Rice	-1.110*	0.212	[-1.526, -0.693]	-1.181*	0.487	[-2.136, -0.226]
Wheat	-1.051*	0.153	[-1.351, -0.751]	-0.977	0.847	[-2.637, 0.682]
Maize	-1.015*	0.243	[-1.490, -0.539]	-0.974	0.727	[-2.399, 0.450]

Estimates significant at the 5% level are indicated thus: '*'. Coeff: Coefficient, Stnd: Standard, Conf: Confidence

It can be seen that only the estimated trend for Rice remains significantly different from zero at the 5% critical level in the difference stationary model. The coefficients for Wheat and Maize, however, would still remain significant at levels of 25.1% and 18.3% respectively. These coefficients would suggest an average annual decline of around one percent *p.a.* in the price of the three cereals relative to the price of manufactured commodities.

[7] Throughout the text, confidence intervals are defined with respect to 'two-tailed' critical values. 90% confidence intervals are listed in appendix III.vii.

Among other food commodities, Palm Oil and Sugar show a significant trend of -1.04% and -1.07% respectively. However, of these two commodities only Sugar is identified as trend stationary by the unit root test and the significant trend does not persist when the model is estimated in first differences. As mentioned above, Banana prices do not show a significant trend in either case.

The estimated trend coefficients for Beef and Lamb are consistently positive, regardless of whether the relative price series is modelled as trend stationary or as difference stationary. The coefficient estimates follow the overall pattern though in so far as they are not significant in the difference stationary model. Trend estimates for food commodities other than cereals or tropical beverages are shown in table 3.2.3.

Table 3.2.3. Trend Estimates for Prices of Other Food Commodities

Commodity	**Trend Coeff.** (*100)	**Stnd. Error** (*100)	95% Conf. Interval	**Drift Coeff.** (*100)	**Stnd. Error** (*100)	95% Conf. Interval
Sugar	-1.067*	0.260	[-1.577, -0.558]	-1.215	1.439	[-4.035, 1.606]
Beef	1.356*	0.591	[0.198, 2.514]	0.807	2.083	[-3.276, 4.889]
Lamb	1.827*	0.358	[1.125, 2.529]	1.548	71.476	[-2.696, 5.792]
Bananas	-0.107	0.311	[-0.717, 0.502]	0.040	0.922	[-1.768, 1.848]
Palm Oil	-1.036*	0.239	[-1.504, -0.567]	-0.713	1.662	[-3.970, 2.544]

Estimates significant at the 5% level are indicated thus: '*'. Coeff: Coefficient, Stnd: Standard, Conf: Confidence

In the case of tropical beverages, there does not appear to be any significant overall trend for the group as a whole. When estimating in levels, the estimate for the coefficient on Tea prices appears to be significant at a 10% significance level. This estimate would suggest a decrease of 0.74% *p.a.* For the remaining tropical beverages, however, trend estimates do not appear to be significant. The closest result to a significant trend estimate for the remaining two commodities would be the estimated trend coefficient on Coffee prices for the trend stationary model. This would suggest an average annual increase of 0.43% at a a significance level of 26.1%. The estimated trends for tropical beverages are given in table 3.2.4.

Table 3.2.4. Trend Estimates for Tropical Beverages

Commodity	**Trend Coeff.** (*100)	**Stnd Error** (*100)	95% Conf. Interval	**Drift Coeff.** (*100)	**Stnd Error** (*100)	95% Conf. Interval
Coffee	0.431	0.381	[-0.328, 1.190]	0.236	2.545	[-4.829, 5.301]
Cocoa	-0.309	0.556	[-1.415, 0.797]	-0.922	2.016	[-4.933, 3.089]
Tea†	-0.744	0.396	[-1.520, 0.032]	-0.954	1.664	[-4.215, 2.307]

Estimates significant at the 5% level are indicated thus: '*'. † Data Series from 1900-1997 only.
Coeff: Coefficient, Stnd: Standard, Conf: Confidence

In the case of agricultural non-food commodities negative and significant trends are present in the trend stationary model in almost all cases (with the exception of Jute and Tobacco where no significant trend is present in either the trend stationary or different stationary model). In most cases, the estimated coefficients for the drift term are clearly insignificantly different from zero in the difference stationary model. The possible exceptions in this case are Wool and Rubber, where significance for the drift term remains at levels of 6.7% and 30% respectively. Table 3.2.5. lists estimated coefficients for the trend of non food agricultural commodities in trend stationary and difference stationary models.

Table 3.2.5. Trend Estimates for Non-food Agricultural Commodities

Commodity	Trend Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval	Drift Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval
Cotton	-0.984*	0.296	[-1.563, -0.404]	-0.786	1.026	[-2.797, 1.226]
Jute	-0.696	0.444	[-1.566, 0.175]	-0.800	1.204	[-3.159, 1.559]
Wool	-1.571*	0.338	[-2.232, -0.909]	-1.450	7.836	[-2.986, 0.086]
Tobacco †	0.468	0.817	[-1.131, 2.069]	0.272	1.436	[-2.543, 3.087]
Rubber	-2.838*	0.420	[-3.660, -2.015]	-3.020	2.900	[-8.704, 2.664]
Timber	1.137*	0.155	[0.833, 1.442]	0.805	1.633	[-2.396, 4.005]

Estimates significant at the 5% level are indicated thus: '*'. † Data Series from 1900-1997 only.
Coeff: Coefficient, Stnd: Standard, Conf: Confidence

For metals finally, there does not appear to be any clear evidence of a common trend for the commodity group. Trend estimates for the relative prices of metals are shown in table 3.2.6.

Table 3.2.6. Trend Estimates for Metals

Commodity	Trend Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval	Drift Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval
Copper	-0.418	0.381	[-1.165, 0.329]	-0.938	1.912	[-4.685, 2.809]
Aluminium	-1.871*	0.208	[-2.279, -1.463]	-1.916*	0.242	[-2.390, -1.443]
Tin	0.096	0.476	[-0.836, 1.029]	-0.330	1.921	[-4.094, 3.434]
Silver	0.022	0.448	[-0.856, 0.901]	-0.331	1.447	[-3.166, 2.505]
Lead	-0.598*	0.276	[-1.139, -0.057]	-0.784	1.895	[-4.498, 2.931]
Zinc	0.059	0.161	[-0.256, 0.374]	0.036	0.198	[-0.352, 0.425]

Estimates significant at the 5% level are indicated thus: '*'. Coeff: Coefficient, Stnd: Standard, Conf: Confidence

Only the estimated trends for Copper, Aluminium and Lead appear to be consistently negative. The trend estimate for Zinc prices is consistently positive though -like the estimated trend for Copper- also consistently insignificant. For the other two consistently insignificant estimated trends (*i.e.* the ones for Tin and Silver) the sign of the trend estimate is actually reversed when the price series are modelled as difference stationary rather than trend stationary. This difference in the sign of point estimates, against the background of the low t-ratios obtained for the coefficient estimates independently of stationarity assumptions, is of course consistent with an insignificant trend coefficient estimate.

It is also worth noting that, although the estimated trend for Zinc is consistently insignificant, the value of the estimated trend coefficient is also rather small, regardless of whether the model is estimated in levels or first differences, and the confidence interval of this estimate is comparatively small. This latter point distinguishes the estimate for Zinc from those for Tin and Silver where, although the estimated coefficient values are small also, the width of their respective confidence intervals indicates that there is still substantial uncertainty surrounding the estimates. In the case of Zinc, one would thus be tempted to infer that the true value of any trend coefficient could indeed be close to zero or at least comparatively small.

3.3. Structural Breaks

When discussing the presence of a secular trend one should pay attention to the issue of structural instability for a number of reasons. When assessing the evolution of a data series over time on the basis of trend estimates, the impression of a persistent trend can easily be produced by structural shifts in particular years: it need not be obvious, when calculating the average annual decline in the relative price of a commodity that part of the overall decline was produced by a single discrete shift rather than a continuous downward trend. On the other hand, a continuous negative trend may be counteracted by a discrete positive shift in a particular year. In graphical plots of the data, the presence of a structural break may be disguised by strong fluctuations of the annual values of the data series around a broken trend. Furthermore, the presence of structural breaks can lead to mistaken inferences on stationarity by reducing the power of unit root tests. In cases, where a single structural break is exogenously inferred, a unit root test which takes account of structural instability has been proposed by Perron (1989). The issue of accounting for endogenous structural breaks in unit root tests is, however, unresolved at present.

In the present case, outliers -and hence possible structural breaks- were inferred endogenously for those years where the residuals from the selected ARIMA models lay outside an interval of ± three sample standard deviations.

What needs to be borne in mind when discussing structural instability in the context of this study is the rather provisional treatment given to it here. In structural models there may be good reasons to model structural breaks separately if particular events are deemed not to be representative of the workings of the causal mechanism under investigation, or where the pattern of the data suggests the possible presence of unrepresentative shocks. There may not be an established consensus as to whether breaks and outliers in structural models should be inferred exogenously or endogenously as a property of the data series itself, but the *a priori* possibility of events that should be discounted from the observed data enjoys wide acceptance.

This is somewhat different in the case of descriptive, univariate data models. Discounting endogenous structural breaks, and multiple structural breaks in particular, corresponds at least partly to a process of adjusting the data to fit the model. For this reason, the treatment of structural breaks and outliers is here confined to the present Chapter and is undertaken only to gain an impression of how far the overall conclusions on trend components and the discrepancies observed over alternative assumptions regarding the order of integration are influenced by the elimination of more extreme variations in the data.

The years with outliers in trend stationary models are presented in table 3.3.1.

Table 3.3.1. Outliers among the Residuals from Trend Stationary Models

Commodity	Years with outliers	Commodity	Years with outliers
Coffee	1976	Jute	1986
Cocoa	1947	Wool	1973
Tea †	1985	Tobacco †	1920, 1960, 1996
Rice	1973	Rubber	1921, 1925, 1950
Wheat	1973	Timber	1993
Maize	1921	Copper	none
Sugar	1963, 1974, 1980	Aluminium	1915
Beef	1915, 1931, 1959	Tin	1986
Lamb	1915, 1931, 1950	Silver	1979, 1980
Banana	none	Lead	1979
Palm Oil	1986	Zinc	1915, 1973
Cotton	none		

† Data series for 1900-1997 only.

Similar tests as for trend stationary models were performed for residuals from difference stationary models selected on the basis of the Schwarz Bayesian

Criterion. The results are given below in table 3.3.2. One should note, moreover, that in table 3.3.2. an outlier corresponding *e.g.* to the year 1901/1902 is entered as 1902.

Table 3.3.2. Outliers among the Residuals from Difference Stationary Models

Commodity	Years with outliers	Commodity	Years with outliers
Coffee	1976	Jute	1986
Cocoa	1947	Wool	1973
Tea†	1985	Tobacco†	1960, 1996
Rice	1973	Rubber	1925, 1950
Wheat	1973	Timber	1993
Maize	1921	Copper	1975
Sugar	1921, 1963, 1974	Aluminium	1915
Beef	1915, 1931, 1959	Tin	1986
Lamb	1915, 1931	Silver	1979
Banana	none	Lead	none
Palm Oil	1986	Zinc	1915, 1973
Cotton	none		

† Data Series from 1900-1997 only.

Information on outliers was then used to decide for which years to include dummy variables for the data series and then re-estimate the models allowing for re-selection using the minimum Schwarz Bayesian Criterion. Dummy variables were specified in such a way as to allow for single additive outliers as well as permanent structural shifts. For single additive outliers, the dummy variable takes the value one in the year in which the outlier occurred and zero in all other years. For structural breaks the dummy takes a value of zero in all years preceding the year for which the outlier was observed and a value of one for all subsequent years. For re-estimation of the difference stationary models, dummies were included in first differences.

3.3.1 The impact on trend estimates

It is to be expected, that unaccounted for structural instability can distort estimates of secular trends by either exaggerating them or -possibly- by dampening them when the unobserved structural shift is counteracting the secular trend. Table 3.3.3. details estimated coefficients for trends from ARIMA models in levels and first differences after structural breaks and single additive outliers have been accounted for. The full details of the estimates for ARIMA models including dummy variables are given in appendix III.iv for the estimates in levels and in appendix III.v. for estimates in first differences.

Table 3.3.3. Estimates of trend and drift coefficients after accounting for outliers

Commodity	Trend coefficient	t-ratio	Coefficient on the drift term	t-ratio
Coffee	-0.006	-0.773	-0.010	-0.402
Cocoa	-0.022	-4.726	-0.021	-1.146
Tea†	-0.000	-0.181	-0.002	-0.112
Rice	-0.021	-3.636	-0.021	-1.112
Wheat	-0.010	-4.227	-0.015	-1.349
Maize	-0.011	-3.259	-0.005	-0.641
Sugar	-0.016	-3.261	-0.010	-0.583
Beef	-0.019	-3.554	-0.016	-0.954
Lamb	0.011	1.601	0.003	0.137
Bananas	-0.001	-0.346	0.000	0.043
Palm Oil	-0.006	-2.930	-0.005	-2.187
Cotton	-0.010	-3.328	-0.008	-0.765
Jute	0.000	0.147	-0.000	-0.021
Wool	-0.013	-3.411	-0.015	-0.826
Tobacco †	-0.003	-0.488	0.000	0.009
Rubber	-0.035	-12.203	-0.051	-1.972
Timber	0.011	6.136	0.005	0.302
Copper	-0.004	-1.097	-0.004	-0.234
Aluminium	-0.026	-5.272	-0.030	-1.514
Tin	0.010	4.045	0.005	0.284
Silver	0.003	0.463	-0.015	-0.854
Lead	-0.001	-0.439	-0.008	-0.414
Zinc	-0.004	-0.952	-0.008	-2.049

† Data Series from 1900-1997 only.

As was the case with estimates of ARIMA models which did not account for structural breaks, there is again a general tendency for t-ratios to decrease noticeably as one moves from trend stationary to difference stationary models. As before, changes in estimates of the trend coefficient tend to be much smaller. Reversals of the sign of the estimated trend coefficient now occur for three commodities (Bananas, Silver and Jute) and for Tobacco. As before, significant trend estimates are given for 13 of the 23 data series, when estimates are made in levels. When estimating in first differences, there are now three price series, with significant drift terms, although the affected price series now are Zinc, Rubber and Palm Oil. In the case of Palm Oil, the t-ratio on the trend coefficient changes remarkably little between the estimates from trend stationary and difference stationary models. In the case of Zinc, the drift term appears significantly different from zero at the 5% significance level when estimates are made in first differences. In contrast to the general pattern observed, however, the estimated coefficient on the trend term for zinc is shown to be statistically insignificant if estimated in levels.

The overall evidence in favour of the presence of secular trends seems not to have changed much. It appears to be worthwhile, however, to compare the impact of structural instability on the trend estimates for individual commodities in several subgroups in detail.

3.3.2. Evaluating the impact of outliers on trend estimates by commodity groups

For the three cereals covered in the data series there still appear to be significant negative trend estimates when estimating in levels, although the estimates now appear insignificant for all estimates in first differences, including Rice. The estimated trends for these commodities are shown in Table 3.3.4.

Table 3.3.4. Trend Estimates for Cereal Prices

Commodity	Trend Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval	Drift Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval
Rice	-2.093*	0.576	[-3.221, -0.965]	-2.136	1.921	[-5.902, 1.630]
Wheat	-0.950*	0.225	[-1.391, -0.510]	-1.468	1.088	[-3.601, 0.665]
Maize	-1.054*	0.323	[-1.688, -0.420]	-0.548	0.856	[-2.225, 1.129]

Estimates significant at the 5% level are indicated thus: '*'. Coeff: Coefficient, Stnd: Standard, Conf: Confidence

The estimated trend coefficients for Rice and Wheat in difference stationary models still would remain significant at the 26.9% and 18.1% significance levels respectively, while the estimated trend coefficient for Maize for the model in first differences appears to be clearly insignificant. For Rice, the absolute value of the estimated negative trend now has increased from around 1% *p.a.* to ca. 2% *p.a.* on average. This can be attributed to the net effect of a positive level shift in 1973 which has been accounted for through the inclusion of a dummy variable. For Wheat, none of the included dummy variables seems to be statistically significant at the 5% level for the model in levels. When re-estimating in first differences, the coefficient for a positive structural break is significant at the 5% level, while the negative coefficient for a single additive outlier is not. Yet, while the estimated downward trend for the model in levels is less pronounced than in the original estimate, the estimated drift coefficient for the model in first differences including dummies actually decreases (*i.e.* it takes a larger negative value). In the case of Maize, the negative impact of a structural break in 1921 appears to be significant at the 10% level when the model is estimated in first differences. This is reflected in a decrease of the estimated average annual decline of the price series from -0.97% *p.a.* to -0.55% *p.a.* When re-estimating the model in levels, the negative coefficient estimate for a single additive outlier appears to be significant, while the positive estimated coefficient on a structural break does not. The estimated downward trend falls slightly from an initial estimate of -1.02% *p.a.* when outliers are not taken into consideration to -1.05% *p.a.* It remains true that the estimated coefficient on the trend term in the trend stationary model is statistically significant while the estimated coefficient on the drift term is not. In so far as the small change in the magnitude of the trend coefficient deserves attention, it appears to contradict the observed change in the estimate of the drift coefficient when dummy variables are included. Given the small magnitude of the change as well as the remaining uncertainty about the significance of the estimates concerned, this result remains difficult to interpret.

Within the overall group of food commodities further consideration may be given to the sub -categories of tropical beverages and other food commodities. Estimates of the trend coefficients for tropical beverages are summarised in table 3.3.5. below.

Table 3.3.5. Trend Estimates for Tropical Beverages

Commodity	Trend Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval	Drift Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval
Coffee	-0.577	0.745	[-2.037, 0.884]	-0.978	2.433	[-5.747, 3.790]
Cocoa	-2.189*	0.463	[-3.097, -1.281]	-2.131	1.859	[-5.775, 1.512]
Tea†	-0.043	0.236	[-0.504, 0.419]	-0.177	1.581	[-3.275, 2.921]

Estimates significant at the 5% level are indicated thus: '*'. † Data Series from 1900-1997 only.
Coeff: Coefficient, Stnd: Standard, Conf: Confidence

In the case of tropical beverages, there now is evidence for a negative trend of around -2% *p.a.* for the price of Cocoa, when the net effect of a positive level shift in 1947 is accounted for. The estimate of the drift term for the difference stationary model is not statistically significant at a level of 5% but would be significant at a significance level of 25.5%. For Coffee, there is still no evidence of a significant trend regardless of whether estimates are made in levels or first differences.

For Tea, the estimated trend coefficient is still insignificant if the model is estimated in either levels or first differences. Compared with the previous estimate -which did not account for dummies- the absolute value of the estimated coefficient is low, and the width of the pertinent confidence interval has decreased. It would thus appear that in this case the incorporation of dummies for structural breaks and outliers strengthens the case for a zero or very small trend coefficient value.

For other food products, outliers and structural breaks also have an impact on the estimated trend coefficient. The coefficient estimates after accounting for the effect of outliers are summarised in table 3.3.6.

Table 3.3.6. Trend Estimates for Other Food Commodities

Commodity	Trend Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval	Drift Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval
Sugar	-1.563*	0.479	[-2.502, -0.623]	-1.020	1.750	[-4.450, 2.409]
Beef	-1.857*	0.522	[-2.880, -0.833]	-1.580	1.656	[-4.825, 1.666]
Lamb	1.144	0.714	[-0.256, 2.543]	0.265	1.938	[-3.534, 4.065]
Bananas	-0.107	0.311	[-0.717, 0.502]	0.040	0.922	[-1.768, 1.848]
Palm Oil	-0.560*	0.191	[-0.934, -0.185]	-0.481*	0.220	[-0.913, -0.050]

Estimates significant at the 5% level are indicated thus: '*'. Coeff: Coefficient, Stnd: Standard, Conf: Confidence

The absolute value of the estimated negative coefficient on the trend for Sugar prices increases in the trend stationary model when the positive net effects of single additive outliers in 1963, 1974 and 1980 are accounted for. Estimating in first differences, the estimate of the downward trend is reduced somewhat from the originally estimated decline of around 1.22% *p.a.* This can be attributed to the combined effects of a large and significant[8] structural break in 1921 and positive breaks in 1963 and 1974.

In the case of Beef, the consideration of structural breaks and single additive outliers appears to have particularly strong effects. The positive trend estimate of around 1.36% *p.a.* which had initially been inferred when estimating in levels is now replaced by a negative trend estimate of -1.86% *p.a.* after the net positive impact of level shifts in 1915, 1931 and 1959 has been accounted for. Estimating

[8] The coefficient on the corresponding dummy is marginally insignificant at the 5% level, with a t-ratio of -1.927, but significant at the 10% level.

in first differences, a similar reversal of coefficient signs is produced by the same sequence of structural breaks.

The residuals for Lamb indicate the presence of outliers in three years, however, only the coefficient estimate for one of them (a structural break in 1915) is shown to be significant when the series is re-estimated in levels including dummies. The impact of the positive breaks reduces the estimate for the positive annual trend coefficient on lamb prices from its original value of 1.83% *p.a.* to around 1.14% *p.a.* The residuals from estimates in first differences show outliers to be present in two years (1915 and 1931) both of which produce evidence of positive structural breaks in the re-estimated model. The effect is again to reduce the value of the estimated positive trend coefficient, which falls from an initial 1.55% *p.a.* to a more modest average annual increase of 0.27%.

The residuals from the estimates on the relative price series for Bananas show no outliers and the corresponding estimated trend coefficients therefore remain unchanged. In the case of Palm Oil, however, a single structural break in 1986 reduces the absolute value of the estimated coefficient for the negative trend from the initial estimates of -1.04% *p.a.* and -0.71% *p.a.*, obtained by estimating in levels and first differences respectively, to the values given in table 3.3.6.

Turning next to non-food agricultural commodities, the results for re-estimated trend coefficients when accounting for outliers are given in table 3.3.7.

Table 3.3.7. Trend Estimates for Non-food Agricultural Commodities

Commodity	Trend Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval	Drift Coeff. (*100)	Stnd. Error (*100)	95% Conf. Interval
Cotton	-0.984*	0.296	[-1.563, -0.404]	-0.786	1.026	[-2.797. 1.226]
Jute	0.045	0.307	[-0.557, 0.648]	-0.044	2.078	[-4.229, 4.140]
Wool	-1.316*	0.386	[-2.072, -0.560]	-1.545	1.870	[-5.210, 2.120]
Tobacco†	-0.303	0.621	[-1.519, 0.914]	0.011	1.228	[-2.396, 2.418]
Rubber	-3.482*	0.285	[-4.042, -2.923]	-5.076	2.574	[-10.120, 0.031]
Timber	1.093*	0.178	[0.744, 1.441]	0.456	1.508	[-2.501, 3.412]

Estimates significant at the 5% level are indicated thus: '*'.
† Data Series from 1900-1997 only.
Coeff: Coefficient, Stnd: Standard, Conf: Confidence

There was no evidence of outliers in the price series for Cotton, so the estimates in this case are unchanged. As for Jute, the original estimate of a significant negative trend now appears to be due to a negative level shift in 1986. Accounting for this structural break, the trend estimates (with a sign reversal between estimates in levels and first differences) now are shown to be highly insignificant. Re-estimating the series for Wool in first differences gives a somewhat higher estimate for the absolute value of the drift term when the effects of a positive single additive outlier and break are taken into consideration. When re-estimating the model in levels, however, the absolute value of the estimated negative trend coefficient is actually reduced. This phenomenon may be the net result of the joint presence of a positive single additive outlier and a negative -albeit insignificant- structural break.

The residuals from the original estimates for the price series of Rubber suggested the existence of outliers in 1921, 1925 and 1950 (when estimated in levels) or 1925 and 1950 when estimated in first differences. Re-estimating in levels, this indicates the presence of a negative structural break in 1921 and of positive structural breaks in 1925 and 1950. Estimates for single additive outliers in these years are positive (though insignificant). The net effect is a larger absolute value for the estimated negative trend coefficient and a markedly higher t-ratio of -12.203. The re-estimated model in first differences shows positive and significant estimates of the coefficients on structural breaks in both years and a positive and a negative estimate for the coefficients on single additive outliers in these years; the two estimated coefficients for single additive outliers are of similar magnitude with values of 0.281 for 1925 and -0.244 for 1950. The net impact on the estimated trend coefficient is an increase in the absolute value of the estimated coefficient when compared with the estimate which did not account for outliers. The estimated coefficient on the drift term now falls from an estimated -0.03 to -0.05 and the P-value for the null hypothesis is now just above 5%, taking a value of 0.516.

Re-estimating the model for Tobacco in levels, the presence of positive structural breaks in 1920 and 1960 combined with a negative structural break in 1996 result in a negative estimate for the trend coefficient. When the model is re-estimated in first differences the presence of a positive structural break in 1960 and a negative structural break in 1996 reduce the positive trend coefficient estimate to a very low value.

Re-estimating the model for Timber shows a structural break in 1993 slightly depressing the value of the estimated positive trend coefficient to a value of 0.0109 from an initial 0.0114 for the estimate in levels. The re-estimated model in first differences shows an estimated drift coefficient of 0.0046 where the original estimate for the drift coefficient was 0.0081.

Considering metals, re-estimation after accounting for the effects of outliers does affect estimation results, but does not provide any more coherent evidence concerning the presence of a trend in this commodity group. Trend coefficient estimates obtained after re-estimating the models for metals are shown in table 3.3.8.

Table 3.3.8. Trend Estimates for Metals

Commodity	**Trend Coefficient** (*100)	**Stnd. Error** (*100)	95% Conf. Interval	**Drift Coefficient** (*100)	**Stnd. Error** (*100)	95% Conf. Interval
Copper	-0.418	0.381	[-1.165, 0.329]	-0.431	1.841	[-4.039, 3.176]
Aluminium	-2.608*	0.495	[-3.577, -1.638]	-2.990	1.974	[-6.860, 0880]
Tin	0.972*	0.240	[0.501, 1.442]	0.504	1.776	[-2.977, 3.985]
Silver	0.257	0.555	[-0.831, 1.344]	-1.492	1.746	[-4.915, 1.931]
Lead	-0.085	0.193	[-0.463, 0.293]	-0.784	1.895	[-4.498, 2.931]
Zinc	-0.409	0.429	[-1.250, 0.432]	-0.849*	0.414	[-1.661, -0.037]

Estimates significant at the 5% level are indicated thus: '*'. Stnd: Standard, Conf: Confidence

Since -when modelling the series as trend stationary- there is no evidence for the presence of outliers in the case of Copper the model has been re-estimated only in first differences. This yields negative coefficient estimates for both a single additive outlier and a negative structural break in 1975. The coefficient on the single additive outlier seems to be clearly insignificant, while the coefficient estimate on structural break is only marginally insignificant at the 5% level -*i.e.* it would be significant at a 5.72% significance level. The estimated coefficient on the drift term nevertheless increases from an original -0.0094 to -0.0043.

Considering the overall positive effect of a structural break in 1915, the estimated trend coefficient for Aluminium decreases from -0.0187 to -0.0261. When the model is re-estimated in levels. Re-estimating the model in first differences, the structural break leads to a fall in the estimated coefficient on the drift term from -0.0192 to -0.0299, although the result also differs from the initial estimate in so far as the estimate is no longer significant at the 5% level but merely would retain significance at a level of 13.3%. In the case of Tin, the presence of a negative structural break now leads to a higher estimated coefficient value for the estimate in levels (*cf.* tables 3.3.8 and 3.2.6.) while for the re-estimated model in first differences the initial negative coefficient estimate on the drift term is now replaced by a positive, though still insignificant, estimated coefficient of 0.005.

Residuals for the initial estimate in levels indicate two outliers for Silver (in 1979 and 1980), while residuals from the corresponding estimate in first differences only show an outlier for 1979. Re-estimation in levels indicates the presence of only one significant coefficient among the included dummies[9]. This identifies a positive single additive outlier in 1979. The value of the estimated trend coefficient increases somewhat but remains insignificant. For the re-estimated model in first differences, there appears to be a significant positive coefficient for a structural break as well as a significant coefficient for the single additive outlier, although the latter now takes a negative sign. The parallel presence of a positive trend

[9] There are insignificant negative coefficients for two structural breaks and a positive insignificant coefficient for a single additive outlier.

coefficient for the trend stationary model and a negative drift coefficient for the difference stationary model are maintained after accounting for outliers with the estimate of the drift coefficient falling further in value to -0.0149 from an initial level of -0.0033.

When the price series for Lead is re-modelled in levels to account for a potential outlier in 1979 the re-estimated equation indicates the presence of a negative and significant coefficient on the structural break dummy as well as a positive and significant coefficient for a single additive outlier. The net result appears to weaken the extent of the previously estimated downward trend (the coefficient increases from -0.006 to -0.001). It should be noted though that the trend coefficient estimate in the re-estimated model is no longer significant. For the model in first differences there were no outliers and no re-estimation has been necessary.

The series for Zinc was re-estimated including dummy variables for outliers for the years 1915 and 1973. When re-estimating in levels, the coefficients on all the dummy variables took positive signs but were insignificant. (Only the coefficient on what could be a structural break in 1915 was marginally insignificant for a 10 percent critical value.) The previous, positive estimate of 0.0006 for the trend coefficient now takes a value of -0.0041 but -given a t-ratio of 0.95- still appears to be insignificant. When re-estimating the model in first differences, the estimated coefficient on the dummy variable suggesting a structural break in 1915 is statistically significant and positive, while the estimated coefficient for the dummy variable indicating a structural break in 1973 is also positive and significant at a 10 percent significance level. As in the trend stationary model, the coefficient estimate for the drift term now takes a negative value. In contrast to the trend stationary model this estimate now appears to be significant, however.

3.3.3. Overall evidence on secular trends

If one were to take the results obtained above without further caution, *i.e.* if one were to decide on the order of integration of the relative price series covered *a priori*, relying on the outcome of unit root tests, then it would seem that a significant trend has been established for four of the five commodities identified as trend stationary. (The commodities in question are Sugar, Lamb, Timber and Aluminium.) Of these, Lamb and Timber have positive and significant trend coefficients. These general results remain mostly unaltered when outliers and structural breaks are taken into consideration though the estimated trend coefficient on Lamb prices now becomes insignificant. Among those commodities, for which the unit root null hypothesis could not be rejected, a significant drift coefficient estimate has been obtained for Rice only, while the one for Wool would be significant at a 10% level. For the remaining price series identified as stationary in first differences (and for Zinc among the series presumed trend stationary), there is no significant trend or drift coefficient estimate at all when structural breaks are not accounted for.

It has been mentioned above that exclusive reliance on unit root tests may well lead to mistaken inference on the order of integration of a data series. It will also be recalled that this is due not only to the possible impact of a large moving average component in the residual process but that similar problems may arise if one fails to account for structural instability. While it has been attempted here to take account of structural breaks under different model alternatives there is no guarantee that the number of structural breaks considered is complete. (On the other hand, the occurrence of insignificant coefficient estimates for some of the included dummy variables, casts doubts on the endogenously inferred outliers and structural breaks.) Furthermore, the incorporation of dummies for possible structural breaks does not resolve the issue of uncertainty about the true order of integration of the series.

It should be obvious then that a test which attempts to establish the order of integration *a priori* leaves a considerable degree of uncertainty in the interpretation of test results. (In the present study, unit root tests have only been conducted without accounting for structural instability. Since structural breaks have been inferred endogenously, and since there is so far no consensus as to how to account for endogenous structural breaks, it appears to be advisable to remain agnostic on the issue for the time being.)

It can be argued that there is a further motive for caution in interpreting the present results since the reliability of t-tests to asses the significance of coefficients depends not only on stationarity but on the presence of normally distributed residuals. Appendix III.vi reports the results of Bowman-Shelton normality tests on the residuals obtained. It can be shown that there are significant departures from normality for some of the price series, although there is some improvement when outliers are taken into account (*cf.* Appendix III.vi for details).

While t-tests may provide information on the statistical significance of positive or negative trend estimates one should be aware of the limitations of such a test statistic. In particular, it is worth noting that by testing whether or not an estimated coefficient is significantly different from zero significance tests are implicitly providing information on the likelihood of the true coefficient having a positive or negative sign. While the question of sign determinacy is crucial in an attempt to infer the presence or absence of a secular trend one should not overlook the more general question of the precision of the point estimate. It is therefore worth considering confidence intervals as well as point estimates and t-ratios.

A careful interpretation of the present results should at least not be conditional upon any pre-established premise on the order of integration of the individual data series. Consideration should therefore be given to estimation results in both levels and first differences when considering the evidence on the presence of a secular trend. On these premises, there seems to be no clear evidence in favour of a generalised negative trend among relative primary commodity prices -neither for individual price series nor for most commodity groups. While consistently negative point estimates are obtained for the trend coefficients for some commodities and for cereals as a group, there is still uncertainty surrounding the significance of these estimates.

3.4. Comparison with Other Studies

A number of other studies have investigated the evidence in favour of a significant negative trend using composite indices of primary commodity prices and often relying on the GYCPI which served as a basis for this study. The original study by

Grilli and Yang (1988) as well as later articles by Cuddington (1992) and León and Soto (1997) looked into the evolution of disaggregated price series. The present section will focus initially on the results for commodity subgroups reported in Grilli and Yang (1988) and then turn to the results reported by Cuddington (1992) and León and Soto (1997) for individual relative price series.

3.4.1. A comparison with the results obtained by Grilli and Yang (1988)

Grilli and Yang estimate trends for four main commodity groups: food prices, non-food agricultural commodities and metal prices. They estimate an ordinary least squares regression of the group index on a constant and trend, modelling all price series as trend stationary[10].

Food prices are reported to follow a negative trend of - 0.36% *p.a.* Within this category a 'strong positive trend' (of 0.63% *p.a.*) is reported for the relative price of tropical beverages. A strict comparison is made difficult by the fact that in the present study, estimates were made for individual commodity price series, while Grilli and Yang only disaggregated to the level of sub-indices of their overall composite index. Moreover, the original data series used by Grilli and Yang only extended as far as 1986, while the present study covers average annual price data up to 1998[11]. However, some informal comparisons should be possible. The estimated trend coefficients from estimates in levels do indeed show a number of significant, negative trend coefficient estimates for relative food prices. These tend to be of a magnitude of around one percent or more in absolute value but are

[10]Grilli and Yang (1988) also report that they used a Maximum Likelihood procedure to correct for serial correlation.

[11]Except for the data series for Tea and Tobacco, where data are available up to 1997 only.

counterbalanced to some degree by positive trend estimates for some commodities (*e.g.* Lamb and Beef). Commenting further on this sub-index, Grilli and Yang also report a 'strong positive trend' for tropical beverages, for which no evidence has been found in this study, regardless of whether the model was estimated in levels or first differences. Rather it appears that the only potentially significant trend estimate in this category (*i.e.* the one for Tea) is negative. Furthermore, when taking structural breaks into consideration, the relative price series for Cocoa appears to follow a significant negative trend when estimated in levels.

The estimate of an average annual decline of about 0.84 percent for non-food agricultural prices still seems to be reasonably similar to the results obtained here, however a negative trend of around 0.82 percent for the relative price of metals appears to be difficult to accept as representative for the results obtained here, given that no clear evidence of a persistent, common trend for the commodities in this group has been obtained. Even when the series are modelled in levels no more than two price series (Aluminium and Lead) show strong evidence in favour of a negative trend.

3.4.2. A comparison with the results obtained by Cuddington (1992)

Cuddington (1992), worked with 26 data series on individual primary commodity prices from the original data set used by Grilli and Yang, including fuel and covering the period 1900-1983. Trend estimates were obtained by first testing for stationarity using unit root tests[12] and then deciding on the order of integration. Subsequently, relative prices of primary commodities were estimated by regression

[12] Cuddington uses the Said Dickey test and Perron tests as appropriate.

on a constant and trend, allowing for an autoregressive moving average process for the residual. If assumptions on the order of integration had been formed exclusively on the basis of unit root tests for those commodities covered by Cuddington (1992) as well as in the present study, then inferences about stationarity would agree with those made by Cuddington in all but eight cases: Aluminium, Coffee, Rice, Wheat, Maize, Palm Oil, Tin and Lead. Aluminium was identified as I(0) in the present study and as I(1) by Cuddington. Coffee, Rice, Wheat, Maize, Palm Oil, Tin and Lead appeared to be I(1) in the present study but are described as I(0) by Cuddington. In the case of Coffee Cuddington (*op. cit.*) includes a dummy for 1950, while in the present case an outlier, identifying a structural break, was found for 1976. Again the scope for comparison is somewhat limited as the data-sets used in both cases are not exactly identical, with the data-set for the present study covering a longer period and being chained to the original GYCPI from 1960. Table 3.4.1. compares estimated trend coefficients from the present study and Cuddington (1992). Comparisons are made for estimates obtained here. Cuddington's inferences on the order of integration are reported together with his trend coefficient estimates in column two of table 3.4.1.

Table 3.4.1. Trend estimates from the present study and Cuddington (1992)

Commodity	Trend (Cuddington)	Trend (no dummies)	Drift (no dummies)	Trend (with dummies)	Drift (with dummies)
Coffee	0.000 (TS)	0.004	0.002	-0.006	-0.010
Cocoa	-0.001 (DS)	-0.003	-0.009	-0.022	-0.021
Tea	-0.005 (DS)	-0.007	-0.010	-0.000	-0.002
Rice	-0.006 (TS)	-0.011	-0.012	-0.021	-0.021
Wheat	-0.007 (TS)	-0.011	-0.010	-0.010	-0.015
Maize	-0.007 (TS)	-0.010	-0.010	-0.011	-0.005
Sugar	-0.007 (TS)	-0.011	-0.012	-0.016	-0.010
Beef	0.019 (DS)	0.014	0.008	-0.019	-0.016
Lamb	0.019 (TS)	0.018	0.015	0.011	0.003
Bananas	0.004 (DS)	-0.001	0.000	-0.001	0.000
Palm Oil	-0.006 (TS)	-0.010	-0.007	-0.006	-0.005
Cotton	-0.002 (DS)	-0.010	-0.008	-0.010	-0.008
Jute	-0.001 (DS)	-0.007	-0.008	0.000	-0.000
Wool	-0.010 (DS)	-0.016	-0.014	-0.013	-0.015
Tobacco	0.015 (DS)	0.005	0.003	-0.003	0.000
Rubber	-0.025 (DS)	-0.028	-0.030	-0.035	-0.051
Timber	0.011 (TS)	0.011	0.008	0.011	0.005
Copper	-0.005 (DS)	-0.004	-0.009	-0.004	-0.004
Aluminium	-0.015 (DS)	-0.019	-0.019	-0.026	-0.030
Tin	0.012 (TS)	0.001	-0.003	0.010	0.005
Silver	0.010 (DS)	0.000	-0.003	0.003	-0.015
Lead	-0.001 (TS)	-0.006	-0.008	-0.001	-0.008
Zinc	0.000 (TS)	0.001	0.000	-0.004	-0.008

TS: trend stationary series, DS: difference stationary series.

There does not appear to be any general pattern in comparative coefficient size suggesting systematic over- or underestimation by one or the other approach. In the majority of cases, the same sign has been obtained for the estimated coefficients. Although there are partial discrepancies[13] in the cases of Bananas, Tin and Silver.

3.4.3. A comparison with the results of León and Soto (1997)

León and Soto (1997) extended the data underlying the Grilli and Yang commodity Price Index up to 1992 and estimate trends for the composite indices covered by

[13] For Coffee, Beef, Jute, Tobacco and Zinc different coefficient signs are obtained only when dummy variables are included in the model.

Grilli and Yang as well as for individual commodity price series. They aim to allow for endogenously inferred structural breaks in Unit Root testing (using the recursive Zivot Andrews test (*cf.* León and Soto (1997, *pp.*353-354)) and classify 20 out of 24 series as trend stationary. The estimates for the price series are trend stationary or difference stationary univariate models allowing for ARMA errors as in the case of Cuddington (1992) and including dummy variables in a number of cases.

As was the case in comparing the results of Cuddington (1992) with the estimates obtained here, there does not appear to be any systematic pattern of differences in the magnitude or the sign of trend or drift coefficients obtained in the study by León and Soto and the coefficient estimates obtained here. Again, the signs of the coefficient estimates obtained agree in most cases. The exceptions now occur for Coffee, Jute and Wool were the trend and drift coefficients take different signs and Tin, Silver and Bananas, where different signs have been obtained for the drift coefficient estimate[14]. The trend or drift coefficient estimates obtained by León and Soto together with the inferred order of integration are reported in Table 3.4.2 below. The trend and drift coefficient estimates obtained for various models in the present study are again given in columns three to six.

Regarding the inferred order of integration of the series, the unit root test employed by León and Soto (1997) identifies a larger number of trend stationary series than either Cuddington (1992) or than has been found here. All the series classified as trend stationary by the ADF test used here are also classified as trend stationary by

[14] Among the models incorporating dummies different coefficient signs to the ones reported by León and Soto (1997) were obtained for Beef and Zinc.

León and Soto. The series modelled as difference stationary by León and Soto (Cocoa, Beef, Bananas and Silver) would also be classified as difference stationary by the ADF test results reported here. In all the remaining cases, the test results of León and Soto indicate trend stationary series, while the ADF test results reported here suggest that the remaining series should be classified as stationary in first differences.

Table 3.4.2. Trend estimates from the present study and León and Soto (1997)

Commodity	Trend (León and Soto)	Trend (no dummies)	Drift (no dummies)	Trend (with dummies)	Drift (with dummies)
Coffee	-0.005 (TS)	0.004	0.002	-0.006	-0.010
Cocoa	-0.550 (DS)	-0.003	-0.009	-0.022	-0.021
Tea	-0.001 (TS)	-0.007	-0.010	-0.000	-0.002
Rice	-0.012 (TS)	-0.011	-0.012	-0.021	-0.021
Wheat	-0.009 (TS)	-0.011	-0.010	-0.010	-0.015
Maize	-0.008 (TS)	-0.010	-0.010	-0.011	-0.005
Sugar	-0.005 (TS)	-0.011	-0.012	-0.016	-0.010
Beef	0.020 (DS)	0.014	0.008	-0.019	-0.016
Lamb	0.017 (TS)	0.018	0.015	0.011	0.003
Bananas	-0.013 (DS)	-0.001	0.000	-0.001	0.000
Palm Oil	-0.006 (TS)	-0.010	-0.007	-0.006	-0.005
Cotton	-0.001 (TS)	-0.010	-0.008	-0.010	-0.008
Jute	0.002 (TS)	-0.007	-0.008	0.000	-0.000
Wool	0.001 (TS)	-0.016	-0.014	-0.013	-0.015
Tobacco	0.004 (TS)	0.005	0.003	-0.003	0.000
Rubber	-0.018 (TS)	-0.028	-0.030	-0.003	-0.051
Timber	0.008 (TS)	0.011	0.008	0.011	0.005
Copper	-0.012 (TS)	-0.004	-0.009	-0.004	-0.004
Aluminium	-0.027 (TS)	-0.019	-0.019	-0.026	-0.030
Tin	0.010 (TS)	0.001	-0.003	0.010	0.005
Silver	-0.002 (DS)	0.000	-0.003	0.003	-0.015
Lead	-0.010 (TS)	-0.006	-0.008	-0.001	-0.008
Zinc	0.000 (TS)	0.001	0.000	-0.004	-0.008

TS: trend stationary series, DS: difference stationary series.

It is worth noting that in the case of Cocoa, where the difference in the magnitude of the trend coefficient reported by León and Soto is large, a constant and trend term were incorporated in the difference stationary model estimated by León and Soto. The estimated coefficient on the constant is -0.200, while the coefficient estimate for the trend term takes the value of -0.550 reported in the table above.

3.5 Conclusion

In this chapter, it was confirmed that conclusions on the statistical significance of trend and drift coefficient estimates and conclusions on the order of integration of the underlying data generating process are interdependent. One should also consider the possibility that this may be so not only in the familiar case where trend stationary models are fitted to integrated series but also in cases where a trend stationary series is wrongly treated as difference stationary.

The outliers observed among the residuals from the estimates and subsequent estimates incorporating dummy variables suggest the possibility of structural breaks in a number of the relative price series. This provides further reasons to be cautious about *a priori* unit root tests, as the results of these can be influenced by structural breaks. However, the incorporation of dummy variables to account for structural breaks does not resolve the existing uncertainty about the presence of trend or drift components.

Against the background of the problems discussed so far it has not been possible to reach a definite conclusion on the presence of trend and drift components in relative primary commodity price series. The following chapter will therefore turn to a more detailed investigation of the evidence in favour of trend components in the presence of sustained uncertainty surrounding the order of integration of the data series.

Appendix III.i. Unit Root Test Results

The Augmented Dickey-Fuller Tests

This appendix gives results for the Augmented Dickey Fuller (ADF) test. The tests were conducted for the data series in levels including up to five lagged differenced terms, a constant and a trend term. Lagged differenced terms were eliminated using general to specific testing, eliminating lagged terms which appeared statistically insignificant at the 5% level and re-estimating one by one. (This assures, that the ADF test reduces to the normal Dickey-Fuller test with constant and trend where none of the lagged terms are significant.) For those commodities, for which the null hypothesis of a unit root could not be rejected in this testing procedure, ADF tests were also conducted for the first differenced series, again employing general to specific testing, and including a constant only in the testing equation. Following Enders (1995), the 5% critical value for the ADF test including constant and trend for 100 observations[1] is ±3.45 the corresponding value for the ADF equation including a constant only is ±2.89.

The results obtained for the ADF test for series in levels are detailed below. The general format for the ADF test equation is:

$$\Delta p_t = \alpha + \beta t + \rho^* p_{t-1} + \sum_{i=1}^{5} \gamma_i \Delta p_{t-i}$$

where standard errors are given in parentheses below the estimated coefficient values, and the t-ratio on the lagged dependent variable (τ) is reported separately. Generally, p_{t-i} is the price of the commodity in question in period t-i, α is the

[1] The series for Tea and Tobacco only contain 98 observations before differencing. The remaining data series have 99 observations. For this sample size and the estimates reported here, the precise critical value should be put at +/- 3.46. This does not affect the conclusions reached.

intercept term, β the trend coefficient and ρ* the coefficient for the Dickey-Fuller test and γ_i the coefficient on the ith lagged differenced term. The trend term is omitted in cases where the estimated coefficient value is zero when rounded to three digits.

Coffee:

$$\Delta p_t = -\underset{(0.080)}{0.180} + \underset{(0.001)}{0.001}\, t - \underset{(0.061)}{0.190}\, p_{t-1}$$

$$\tau = -3.121$$

Cocoa:

$$\Delta p_t = -\underset{(0.075)}{0.137} - \underset{(0.053)}{0.114}\, p_{t-1} + \underset{(0.100)}{0.135}\, \Delta p_{t-1} - \underset{(0.101)}{0.256}\, \Delta p_{t-2}$$

$$\tau = -2.158$$

Tea:

$$\Delta p_t = \underset{(0.036)}{0.039} - \underset{(0.001)}{0.001}\, t - \underset{(0.048)}{0.108}\, p_{t-1}$$

$$\tau = -2.248$$

Rice:

$$\Delta p_t = \underset{(0.061)}{0.170} - \underset{(0.001)}{0.003}\, t - \underset{(0.068)}{0.224}\, p_{t-1} + \underset{(0.097)}{0.337}\, \Delta p_{t-1} - \underset{(0.102)}{0.210}\, \Delta p_{t-2}$$

$$\tau = -3.271$$

Wheat:

$$\Delta p_t = \underset{(0.083)}{0.220} - \underset{(0.001)}{0.003}\, t - \underset{(0.090)}{0.247}\, p_{t-1} + \underset{(0.107)}{0.243}\, \Delta p_{t-1} - \underset{(0.109)}{0.141}\, \Delta p_{t-2}$$
$$+ \underset{(0.103)}{0.128}\, \Delta p_{t-3} - \underset{(0.105)}{0.262}\, \Delta p_{t-4}$$

$$\tau = -2.740$$

Maize:

$$\Delta p_t = \underset{(0.095)}{0.238} - \underset{(0.001)}{0.003}\, t - \underset{(0.097)}{0.239}\, p_{t-1} + \underset{(0.117)}{0.003}\, \Delta p_{t-1} - \underset{(0.114)}{0.217}\, \Delta p_{t-2}$$
$$- \underset{(0.106)}{0.056}\, \Delta p_{t-3} - \underset{(0.103)}{0.226}\, \Delta p_{t-4}$$

$$\tau = -2.468$$

Sugar:

$$\Delta p_t = \underset{(0.087)}{0.244} - \underset{(0.001)}{0.004}\, t - \underset{(0.077)}{0.339}\, p_{t-1}$$

$$\tau = -4.391$$

Beef:

$$\Delta p_t = -\underset{(0.089)}{0.119} + \underset{(0.001)}{0.001}\, t - \underset{(0.045)}{0.085}\, p_{t-1}$$

$$\tau = -1.876$$

Lamb:

$$\Delta p_t = -\underset{(0.105)}{0.312} + \underset{(0.001)}{0.003}\, t - \underset{(0.051)}{0.181}\, p_{t-1} + \underset{(0.099)}{0.087}\, \Delta p_{t-1} + \underset{(0.099)}{0.058}\, \Delta p_{t-2}$$
$$+ \underset{(0.099)}{0.092}\, \Delta p_{t-3} + \underset{(0.099)}{0.426}\, \Delta p_{t-4}$$

$$\tau = -3.535$$

Bananas:

$$\Delta p_t = \underset{(0.023)}{0.054} - \underset{(0.000)}{0.001}\, t - \underset{(0.040)}{0.102}\, p_{t-1}$$

$$\tau = -2.521$$

Palm Oil:

$$\Delta p_t = \underset{(0.062)}{0.134} - \underset{(0.001)}{0.003}\, t - \underset{(0.075)}{0.224}\, p_{t-1} + \underset{(0.100)}{0.155}\, \Delta p_{t-1} - \underset{(0.101)}{0.255}\, \Delta p_{t-2}$$

$$\tau = -2.991$$

Cotton:

$$\Delta p_t = \underset{(0.058)}{0.146} - \underset{(0.001)}{0.002}\, t - \underset{(0.065)}{0.166}\, p_{t-1} + \underset{(0.100)}{0.073}\, \Delta p_{t-1} - \underset{(0.100)}{0.264}\, \Delta p_{t-2}$$

$$\tau = -2.553$$

Jute:

$$\Delta p_t = \underset{(0.054)}{0.104} - \underset{(0.001)}{0.002}\, t - \underset{(0.063)}{0.127}\, p_{t-1} + \underset{(0.103)}{0.009}\, \Delta p_{t-1} - \underset{(0.102)}{0.267}\, \Delta p_{t-2}$$

$$\tau = -2.027$$

Wool:

$$\Delta p_t = \underset{(0.108)}{0.251} - \underset{(0.001)}{0.004}\, t - \underset{(0.073)}{0.155}\, p_{t-1} - \underset{(0.109)}{0.078}\, \Delta p_{t-1} - \underset{(0.108)}{0.313}\, \Delta p_{t-2}$$
$$- \underset{(0.103)}{0.082}\, \Delta p_{t-3} - \underset{(0.103)}{0.252}\, \Delta p_{t-4}$$

$$\tau = -2.137$$

Tobacco:

$$\Delta p_t = \underset{(0.050)}{0.054} - \underset{(0.001)}{0.001}\, t - \underset{(0.051)}{0.022}\, p_{t-1} + \underset{(0.112)}{0.017}\, \Delta p_{t-1} + \underset{(0.112)}{0.009}\, \Delta p_{t-2}$$
$$- \underset{(0.111)}{0.050}\, \Delta p_{t-3} - \underset{(0.111)}{0.259}\, \Delta p_{t-4}$$

$$\tau = -0.437$$

Rubber:

$$\Delta p_t = \underset{(0.132)}{0.357} - \underset{(0.002)}{0.005}\, t - \underset{(0.061)}{0.197}\, p_{t-1}$$

$$\tau = -3.217$$

Timber:

$$\Delta p_t = -\underset{(0.110)}{0.383} + \underset{(0.001)}{0.004}\, t - \underset{(0.101)}{0.395}\, p_{t-1} + \underset{(0.118)}{0.155}\, \Delta p_{t-1} - \underset{(0.111)}{0.005}\, \Delta p_{t-2} + \underset{(0.107)}{0.216}\, \Delta p_{t-3}$$

$$\tau = -3.925$$

Copper:

$$\Delta p_t = \underset{(0.038)}{0.017} - \underset{(0.054)}{0.139}\, p_{t-1}$$

$$\tau = -2.600$$

Aluminium:

$$\Delta p_t = \underset{(0.090)}{0.324} - \underset{(0.001)}{0.005}\, t - \underset{(0.062)}{0.255}\, p_{t-1} + \underset{(0.099)}{0.301}\, \Delta p_{t-1}$$

$$\tau = -4.128$$

Tin:

$$\Delta p_t = -\underset{(0.064)}{0.081} - \underset{(0.050)}{0.103}\, p_{t-1}$$

$$\tau = -2.062$$

Silver:

$$\Delta p_t = -\underset{(0.058)}{0.087} - \underset{(0.049)}{0.092}\, p_{t-1} + \underset{(0.101)}{0.087}\, \Delta p_{t-1} - \underset{(0.101)}{0.265}\, \Delta p_{t-2}$$

$$\tau = -1.866$$

Lead:

$$\Delta p_t = \underset{(0.037)}{0.001} - \underset{(0.001)}{0.001}\, t - \underset{(0.063)}{0.200}\, p_{t-1}$$

$$\tau = -3.169$$

Zinc:

$$\Delta p_t = \underset{(0.041)}{0.005} - \underset{(0.086)}{0.437}\, p_{t-1} + \underset{(0.101)}{0.221}\, \Delta p_{t-1}$$

$$\tau = -5.106$$

For those series which could not be identified as trend stationary, further unit root tests were conducted for the series in first differences, to confirm that the first differenced series are indeed stationary. The general format for the ADF test for the differenced data series was as follows:

$$\Delta^2 p_t = a + \rho^* \Delta p_{t-1} + \sum_{i=1}^{5} \gamma_i \Delta^2 p_{t-i}$$

where Δ^2 indicates a second difference, and the remaining elements of the equation are as before. The results obtained were as follows:

Coffee:

$$\Delta^2 p_t = \underset{(0.025)}{0.007} - \underset{(0.148)}{1.313}\, \Delta p_{t-1} + \underset{(0.103)}{0.233}\, \Delta^2 p_{t-1}$$

$\tau = -8.897$

Cocoa:

$$\Delta^2 p_t = -\underset{(0.025)}{0.012} - \underset{(0.135)}{1.234}\, \Delta p_{t-1} + \underset{(0.099)}{0.317}\, \Delta^2 p_{t-1}$$

$\tau = -9.146$

Tea:

$$\Delta^2 p_t = -\underset{(0.017)}{0.014} - \underset{(0.145)}{1.240}\, \Delta p_{t-1} + \underset{(0.103)}{0.258}\, \Delta^2 p_{t-1}$$

$\tau = -8.557$

Rice:

$$\Delta^2 p_t = -\underset{(0.017)}{0.014} - \underset{(0.202)}{1.393}\, \Delta p_{t-1} + \underset{(0.164)}{0.632}\, \Delta^2 p_{t-1} + \underset{(0.128)}{0.229}\, \Delta^2 p_{t-2} + \underset{(0.102)}{0.221}\, \Delta^2 p_{t-3}$$

$\tau = -6.893$

Wheat:

$$\Delta^2 p_t = -\underset{(0.017)}{0.020} - \underset{(0.297)}{1.950}\,\Delta p_{t-1} + \underset{(0.256)}{1.040}\,\Delta^2 p_{t-1} + \underset{(0.213)}{0.672}\,\Delta^2 p_{t-2} + \underset{(0.182)}{0.711}\,\Delta^2 p_{t-3}$$
$$+\underset{(0.141)}{0.290}\,\Delta^2 p_{t-4} + \underset{(0.105)}{0.276}\,\Delta^2 p_{t-5}$$

$\tau = -6.566$

Maize:

$$\Delta^2 p_t = -\underset{(0.021)}{0.019} - \underset{(0.247)}{1.913}\,\Delta p_{t-1} + \underset{(0.202)}{0.773}\,\Delta^2 p_{t-1} + \underset{(0.150)}{0.423}\,\Delta^2 p_{t-2} + \underset{(0.101)}{0.284}\,\Delta^2 p_{t-3}$$

$\tau = -7.739$

Beef:

$$\Delta^2 p_t = \underset{(0.021)}{0.007} - \underset{(0.102)}{0.953}\,\Delta p_{t-1}$$

$\tau = -9.297$

Bananas:

$$\Delta^2 p_t = -\underset{(0.009)}{0.001} - \underset{(0.102)}{0.939}\,\Delta p_{t-1}$$

$\tau = -9.208$

Palm Oil:

$$\Delta^2 p_t = -\underset{(0.023)}{0.014} - \underset{(0.283)}{1.686}\,\Delta p_{t-1} + \underset{(0.244)}{0.724}\,\Delta^2 p_{t-1} + \underset{(0.204)}{0.334}\,\Delta^2 p_{t-2}$$
$$+\underset{(0.147)}{0.260}\,\Delta^2 p_{t-3} + \underset{(0.107)}{0.210}\,\Delta^2 p_{t-4}$$

$\tau = -5.961$

Cotton:

$$\Delta^2 p_t = -\underset{(0.016)}{0.012} - \underset{(0.138)}{1.325}\,\Delta p_{t-1} + \underset{(0.098)}{0.330}\,\Delta^2 p_{t-1}$$

$\tau = -9.575$

Jute:

$$\Delta^2 p_t = -\underset{(0.023)}{0.013} - \underset{(0.238)}{1.552}\,\Delta p_{t-1} + \underset{(0.201)}{0.547}\,\Delta^2 p_{t-1} + \underset{(0.149)}{0.164}\,\Delta^2 p_{t-2} + \underset{(0.105)}{0.236}\,\Delta^2 p_{t-3}$$

$\tau = -6.522$

Wool:

$$\Delta^2 p_t = -\underset{(0.020)}{0.028} - \underset{(0.253)}{1.884}\,\Delta p_{t-1} + \underset{(0.208)}{0.737}\,\Delta^2 p_{t-1} + \underset{(0.154)}{0.369}\,\Delta^2 p_{t-2} + \underset{(0.103)}{0.267}\,\Delta^2 p_{t-3}$$

$\tau = -7.451$

Tobacco:

$$\Delta^2 p_t = \underset{(0.015)}{0.003} - \underset{(0.205)}{1.186}\,\Delta p_{t-1} + \underset{(0.179)}{0.231}\,\Delta^2 p_{t-1} + \underset{(0.149)}{0.261}\,\Delta^2 p_{t-2} + \underset{(0.108)}{0.233}\,\Delta^2 p_{t-3}$$

$\tau = -5.798$

Rubber:

$$\Delta^2 p_t = -\underset{(0.030)}{0.029} - \underset{(0.103)}{0.958}\,\Delta p_{t-1}$$

$\tau = -9.304$

Copper:

$$\Delta^2 p_t = -\underset{(0.019)}{0.007} - \underset{(0.139)}{1.143}\,\Delta p_t + \underset{(0.101)}{0.207}\,\Delta^2 p_{t-1}$$

$\tau = -8.241$

Tin:

$$\Delta^2 p_t = \underset{(0.019)}{0.002} - \underset{(0.098)}{0.944}\,\Delta p_{t-1}$$

$\tau = -9.609$

Silver:

$$\Delta^2 p_t = -\underset{(0.019)}{0.004} - \underset{(0.137)}{1.270}\,\Delta p_{t-1} + \underset{(0.099)}{0.313}\,\Delta^2 p_{t-1}$$

$\tau = -9.241$

Lead:

$$\Delta^2 p_t = -\underset{(0.019)}{0.011} - \underset{(0.184)}{1.316}\,\Delta p_t + \underset{(0.146)}{0.292}\,\Delta^2 p_{t-1} + \underset{(0.104)}{0.205}\,\Delta^2 p_{t-2}$$

$$\tau = -7.143$$

Appendix III.ii. Estimation Results for Relative Primary Product Price Series in Levels -minimum SBC specifications

This appendix gives details of the estimation results for relative primary product price series. The price series are natural logarithms of the primary commodity price series shown relative to the Manufacturing Unit Value (MUV) index. Price series were modelled for a constant and a linear trend allowing for a general ARMA specification for the error term, *i.e.*

$$p_t = a + \beta t + u_t$$

and

$$u_t - \phi_1 u_{t-1} - \ldots - \phi_p u_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

where p_t is the relevant price series, α the coefficient on the constant, β the coefficient on the linear trend term, ϕ_i the coefficient on the relevant autoregressive error term (u_{t-i}), θ_i the coefficient on the i[th] moving average term (ε_{t-i}). The subscript t indicates time period t and t-i the i[th] lag. ARMA models were estimated for all ARIMA(p,0,q) specifications such that p+q≤5 and the preferred model was identified by SBC. The estimation results were obtained using the ARIMA.SRC procedure in GAUSS, and are listed below.

The results shown are the estimated models (with standard errors shown in parentheses) as well as the corresponding value for the Schwarz Bayesian Criterion, and the Ljung Box Q statistic[1] for 12 autocorrelations. P-values for

[1] Johnston and Dinardo (1997) argue that the Ljung Box Q statistic is more appropriate than the Box Pierce statistic as a measure of serial correlation in finite samples. However, they also point out that this test statistic is defined for pure ARMA processes. Given that the present model contains a trend, some caution is in order in interpreting the values obtained.

the Ljung Box Q statistic are given in parentheses after the value of the test statistic. (On this basis and adopting a threshold value of 5% for the test there would be evidence of autocorrelation for the residuals for Cotton and Jute.)

Commodity: Coffee
Model: ARIMA(1,0,0), with constant and trend
SBC = 11.208, **Ljung-Box Q(12):** 9.168 (0.606)
Degrees of freedom: 96

$$p_t = -\underset{(0.224)}{1.029} + \underset{(0.004)}{0.004}t + u_t$$
$$u_t - \underset{(0.060)}{0.803}u_{t-1} = \underset{(0.241)}{\varepsilon_t}$$

Commodity: Cocoa
Model: ARIMA(1,0,0), with constant and trend
SBC = 17.961, **Ljung-Box Q(12):**16.716 (0.117)
Degrees of freedom: 96

$$p_t = -\underset{(0.331)}{0.848} - \underset{(0.006)}{0.003}t + u_t$$
$$u_t - \underset{(0.049)}{0.870}u_{t-1} = \underset{(0.249)}{\varepsilon_t}$$

Commodity: Tea
Model: ARIMA(1,0,0), with constant and trend
SBC = -68.336 , **Ljung-Box Q(12):**13.624 (0.255)
Degrees of freedom: 95

$$p_t = \underset{(0.233)}{0.354} - \underset{(0.004)}{0.007}t + u_t$$
$$u_t - \underset{(0.048)}{0.883}u_{t-1} = \underset{(0.160)}{\varepsilon_t}$$

Commodity: Rice
Model: ARIMA(1,0,1), with constant and trend
SBC = -64.827 , **Ljung-Box Q(12):**7.576 (0.670)
Degrees of freedom: 95

$$p_t = \underset{(0.124)}{0.700} - \underset{(0.002)}{0.011}t + u_t$$
$$u_t - \underset{(0.092)}{0.609}u_{t-1} = \underset{(0.161)}{\varepsilon_t} + \underset{(0.098)}{0.563}\varepsilon_{t-1}$$

Commodity: Wheat
Model: ARIMA(0,0,3), with constant and trend
SBC = -73.142 , **Ljung-Box Q(12):** 6.647 (0.674)
Degrees of freedom: 94

$$p_t = \underset{(0.089)}{0.794} - \underset{(0.002)}{0.011}t + u_t$$
$$u_t = \underset{(0.152)}{\varepsilon_t} + \underset{(0.096)}{1.030}\varepsilon_{t-1} + \underset{(0.135)}{0.566}\varepsilon_{t-2} + \underset{(0.099)}{0.394}\varepsilon_{t-3}$$

Commodity: Maize
Model: ARIMA(1,0,0) with constant and trend
SBC = -17.792 , **Ljung-Box Q(12):**13.534 (0.260)
Degrees of freedom: 96

$$p_t = \underset{(0.141)}{0.785} - \underset{(0.002)}{0.010}t + u_t$$
$$u_t - \underset{(0.071)}{0.720}u_{t-1} = \underset{(0.209)}{\varepsilon_t}$$

Commodity: Sugar
Model: ARIMA(1,0,1) with constant and trend
SBC = 64.951 , **Ljung-Box Q(12):**11.939 (0.289)
Degrees of freedom: 95

$$p_t = \underset{(0.150)}{0.761} - \underset{(0.003)}{0.011}t + u_t$$
$$u_t - \underset{(0.130)}{0.425}u_{t-1} = \underset{(0.311)}{\varepsilon_t} + \underset{(0.132)}{0.413}\varepsilon_{t-1}$$

Commodity: Beef
Model: ARIMA(1,0,0), with constant and trend
SBC = -22.121 , **Ljung-Box Q(12):**10.679 (0.471)
Degrees of freedom: 96

$$p_t = -\underset{(0.350)}{1.602} + \underset{(0.006)}{0.014}t + u_t$$
$$u_t - \underset{(0.044)}{0.905}u_{t-1} = \underset{(0.203)}{\varepsilon_t}$$

Commodity: Lamb
Model: ARIMA(5,0,0), with constant and trend
SBC = -16.512 , **Ljung-Box Q(12):** 3.376 (0.848)
Degrees of freedom: 92

$$p_t = -\underset{(0.207)}{1.779} + \underset{(0.004)}{0.018}t + u_t$$
$$u_t - \underset{(0.094)}{0.897}u_{t-1} + \underset{(0.128)}{0.028}u_{t-2} - \underset{(0.129)}{0.033}u_{t-3} - \underset{(0.130)}{0.321}u_{t-4} + \underset{(0.096)}{0.406}u_{t-5} = \underset{(0.194)}{\varepsilon_t}$$

Commodity: Bananas
Model: ARIMA(1,0,0) with constant and trend
SBC = -181.839 , **Ljung-Box Q(12):**8.438 (0.674)
Degrees of freedom: 96

$$p_t = \underset{(0.189)}{0.207} - \underset{(0.003)}{0.001}t + u_t$$
$$u_t - \underset{(0.037)}{0.926}u_{t-1} = \underset{(0.091)}{\varepsilon_t}$$

Commodity: Palm Oil
Model: ARIMA(1,0,1) with constant and trend
SBC = -13.339 , **Ljung-Box Q(12):**11.709 (0.305)
Degrees of freedom: 95

$$p_t = \underset{(0.139)}{0.516} - \underset{(0.002)}{0.010}t + u_t$$
$$u_t - \underset{(0.102)}{0.591}u_{t-1} = \underset{(0.209)}{\varepsilon_t} + \underset{(0.117)}{0.405}\varepsilon_{t-1}$$

Commodity: Cotton
Model: ARIMA(1,0,0) with constant and trend
SBC = -67.885 , **Ljung-Box Q(12):**20.823 (0.035)
Degrees of freedom: 96

$$p_t = \underset{(0.173)}{0.647} - \underset{(0.003)}{0.010}t + u_t$$
$$u_t - \underset{(0.057)}{0.833}u_{t-1} = \underset{(0.162)}{\varepsilon_t}$$

Commodity: Jute
Model: ARIMA(1,0,0) with constant and trend
SBC = -5.261 , **Ljung-Box Q(12):**20.732 (0.036)
Degrees of freedom: 96

$$p_t = \underset{(0.259)}{0.428} - \underset{(0.004)}{0.007}t + u_t$$
$$u_t - \underset{(0.057)}{0.848}u_{t-1} = \underset{(0.222)}{\varepsilon_t}$$

Commodity: Wool
Model: ARIMA(1,0,0) with constant and trend
SBC = -33.017 , **Ljung-Box Q(12):**7.895 (0.723)
Degrees of freedom: 96

$$p_t = \underset{(0.198)}{1.234} - \underset{(0.003)}{0.016}t + u_t$$
$$u_t - \underset{(0.058)}{0.824}u_{t-1} = \underset{(0.193)}{\varepsilon_t}$$

Commodity: Tobacco
Model: ARIMA(1,0,0) with constant and trend
SBC = -92.171 , **Ljung-Box Q(12):** 11.081 (0.436)
Degrees of freedom: 95

$$p_t = -\underset{(0.439)}{0.713} + \underset{(0.008)}{0.005}t + u_t$$
$$u_t - \underset{(0.041)}{0.953}u_{t-1} = \underset{(0.141)}{\varepsilon_t}$$

Commodity: Rubber
Model: ARIMA(1,0,0) with constant and trend
SBC = 36.437 , **Ljung-Box Q(12):** 9.500 (0.576)
Degrees of freedom: 96

$$p_t = \underset{(0.246)}{1.986} - \underset{(0.004)}{0.028}t + u_t$$
$$u_t - \underset{(0.061)}{0.796}u_{t-1} = \underset{(0.274)}{\varepsilon_t}$$

Commodity: Timber
Model: ARIMA(1,0,0) with constant and trend
SBC = -83.095 , **Ljung-Box Q(12):** 8.208 (0.695)
Degrees of freedom: 96

$$p_t = -\underset{(0.090)}{1.006} + \underset{(0.002)}{0.011}t + u_t$$
$$u_t - \underset{(0.078)}{0.680}u_{t-1} = \underset{(0.150)}{\varepsilon_t}$$

Commodity: Copper
Model: ARIMA(1,0,0) with constant and trend
SBC = -42.081 , **Ljung-Box Q(12):** 6.542 (0.835)
Degrees of freedom: 96

$$p_t = \underset{(0.225)}{0.243} - \underset{(0.004)}{0.004}t + u_t$$
$$u_t - \underset{(0.052)}{0.856}u_{t-1} = \underset{(0.184)}{\varepsilon_t}$$

Commodity: Aluminium
Model: ARIMA(1,0,1) with constant and trend
SBC = -78.002 , **Ljung-Box Q(12):** 2.760 (0.987)
Degrees of freedom: 95

$$p_t = \underset{(0.121)}{1.356} - \underset{(0.002)}{0.019}t + u_t$$
$$u_t - \underset{(0.090)}{0.650}u_{t-1} = \underset{(0.151)}{\varepsilon_t} + \underset{(0.107)}{0.472}\varepsilon_{t-1}$$

Commodity: Tin
Model: ARIMA(1,0,0) with constant and trend
SBC = -39.073 , **Ljung-Box Q(12):** 4.271 (0.961)
Degrees of freedom: 96

$$p_t = -\underset{(0.279)}{0.944} + \underset{(0.005)}{0.001}t + u_t$$
$$u_t - \underset{(0.050)}{0.886}u_{t-1} = \underset{(0.187)}{\varepsilon_t}$$

Commodity: Silver
Model: ARIMA(1,0,0) with constant and trend
SBC = -41.772 , **Ljung-Box Q(12):** 14.702 (0.197)
Degrees of freedom: 96

$$p_t = -\underset{(0.268)}{0.749} + \underset{(0.004)}{0.000}t + u_t$$
$$u_t - \underset{(0.046)}{0.884}u_{t-1} = \underset{(0.184)}{\varepsilon_t}$$

Commodity: Lead
Model: ARIMA(1,0,0) with constant and trend
SBC = -47.206 , **Ljung-Box Q(12):**9.140 (0.609)
Degrees of freedom: 96

$$p_t = -\underset{(0.161)}{0.047} - \underset{(0.003)}{0.006}t + u_t$$
$$u_t - \underset{(0.063)}{0.795}u_{t-1} = \underset{(0.180)}{\varepsilon_t}$$

Commodity: Zinc
Model: ARIMA(1,0,1) with constant and trend
SBC = -26.825 , **Ljung-Box Q(12):**6.370 (0.783)
Degrees of freedom: 95

$$p_t = -\underset{(0.093)}{0.009} + \underset{(0.002)}{0.001}t + u_t$$
$$u_t - \underset{(0.134)}{0.427} = \underset{(0.196)}{\varepsilon_t} + \underset{(0.139)}{0.381}\varepsilon_{t-1}$$

Appendix III.iii. Estimation Results for Relative Primary Product Price Series in First Differences -minimum SBC specifications

This appendix gives details of the estimation results for relative primary product price series. The price series are natural logarithms of the primary commodity price series shown relative to the Manufacturing Unit Value (MUV) index. Price series were modelled including a constant and allowing for a general ARMA specification for the error term, *i.e.*

$$\Delta p_t = \beta + v_t$$

and

$$v_t - \phi_1 v_{t-1} - \ldots - \phi_p v_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

where Δp_t is the relevant price series in first differences, β the coefficient on the constant (*i.e.* the drift term), ϕ_i the coefficient on the relevant autoregressive error term (v_{t-i}), θ_i the coefficient on the i^{th} moving average term (ε_{t-i}). The subscript t indicates time period t and t-i the i^{th} lag. ARIMA models were estimated for all ARIMA(p,1,q) specifications such that $p+q\leq5$ and the preferred model was identified by SBC. The estimation results were obtained using the ARIMA.SRC procedure in GAUSS, and are listed below. Ljung Box Q statistics for 12 autocorrelations are again reported with P-values in parentheses. (At a 5% significance level this would indicate autocorrelation for Wheat and Lamb, although this is not the case for the more highly parameterised difference stationary models presented below in Appendix IV.i for Chapter 4.)

Commodity: Coffee
Model: ARIMA(0,1,0) with constant
SBC = 11.518 , **Ljung-Box Q(12):**11.933 (0.451)
Degrees of freedom: 97

$$\Delta p_t = \underset{(0.025)}{0.002} + v_t$$
$$v_t = \underset{(0.252)}{\varepsilon_t}$$

Commodity: Cocoa
Model: ARIMA(2,1,0) with constant
SBC = 12.852 , **Ljung-Box Q(12):**3.927 (0.916)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.020)}{0.009} + v_t$$
$$v_t - \underset{(0.097)}{0.082} v_{t-1} + \underset{(0.097)}{0.311} v_{t-2} = \underset{(0.244)}{\varepsilon_t}$$

Commodity: Tea
Model: ARIMA(0,1,0) with constant
SBC = -72.034 , **Ljung-Box Q(12):**17.578 (0.129)
Degrees of freedom: 96

$$\Delta p_t = -\underset{(0.017)}{0.010} + v_t$$
$$v_t = \underset{(0.164)}{\varepsilon_t}$$

Commodity: Rice
Model: ARIMA(1,1,2) with constant
SBC = -61.045 , **Ljung-Box Q(12):**8.037 (0.530)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.005)}{0.012} + v_t$$
$$v_t - \underset{(0.140)}{0.551} v_{t-1} = \underset{(0.164)}{\varepsilon_t} - \underset{(0.131)}{0.349} \varepsilon_{t-1} - \underset{(0.093)}{0.542} \varepsilon_{t-2}$$

Commodity: Wheat
Model: ARIMA(0,1,2) with constant
SBC = -71.394 , **Ljung-Box Q(12):**21.786 (0.016)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.008)}{0.010} + v_t$$
$$v_t = \underset{(0.158)}{\varepsilon_t} + \underset{(0.084)}{0.102} \varepsilon_{t-1} - \underset{(0.085)}{0.582} \varepsilon_{t-2}$$

Commodity: Maize
Model: ARIMA(0,1,2) with constant
SBC = -21.473 , **Ljung-Box Q(12):**11.805 (0.298)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.007)}{0.010} + v_t$$
$$v_t = \underset{(0.205)}{\varepsilon_t} - \underset{(0.093)}{0.218}\varepsilon_{t-1} - \underset{(0.093)}{0.441}\varepsilon_{t-2}$$

Commodity: Sugar
Model: ARIMA(0,1,2) with constant
SBC = 68.118 , **Ljung-Box Q(12):**10.349 (0.410)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.014)}{0.012} + v_t$$
$$v_t = \underset{(0.324)}{\varepsilon_t} - \underset{(0.093)}{0.148}\varepsilon_{t-1} - \underset{(0.093)}{0.422}\varepsilon_{t-2}$$

Commodity: Beef
Model: ARIMA(0,1,0) with constant
SBC = -27.798 , **Ljung-Box Q(12):**10.553 (0.568)
Degrees of freedom: 97

$$\Delta p_t = \underset{(0.021)}{0.008} + v_t$$
$$v_t = \underset{(0.206)}{\varepsilon_t}$$

Commodity: Lamb
Model: ARIMA(0,1,0) with constant
SBC = -20.161 , **Ljung-Box Q(12):**23.390 (0.025)
Degrees of freedom: 97

$$\Delta p_t = \underset{(0.022)}{0.015} + v_t$$
$$v_t = \underset{(0.214)}{\varepsilon_t}$$

Commodity: Bananas
Model: ARIMA(0,1,0) with constant
SBC = -187.437 , **Ljung-Box Q(12):**10.778 (0.548)
Degrees of freedom: 97

$$\Delta p_t = \underset{(0.009)}{0.000} + v_t$$
$$v_t = \underset{(0.091)}{\varepsilon_t}$$

Commodity: Palm Oil
Model: ARIMA(2,1,0) with constant
SBC = -12.793 , **Ljung-Box Q(12):** 7.186 (0.618)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.017)}{0.007} + v_t$$

$$v_t - \underset{(0.096)}{0.052} v_{t-1} + \underset{(0.096)}{0.360} v_{t-2} = \underset{(0.214)}{\varepsilon_t}$$

Commodity: Cotton
Model: ARIMA(2,1,2) with constant
SBC = -73.914 , **Ljung-Box Q(12):** 9.176 (0.328)
Degrees of freedom: 93

$$\Delta p_t = -\underset{(0.010)}{0.008} + v_t$$

$$v_t - \underset{(0.077)}{1.303} v_{t-1} + \underset{(0.076)}{0.758} v_{t-2} = \underset{(0.149)}{\varepsilon_t} - \underset{(0.061)}{1.649} \varepsilon_{t-1} + \underset{(0.065)}{0.964} \varepsilon_{t-2}$$

Commodity: Jute
Model: ARIMA(0,1,2) with constant
SBC = -12.668 , **Ljung-Box Q(12):**13.184 (0.214)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.012)}{0.008} + v_t$$

$$v_t = \underset{(0.214)}{\varepsilon_t} - \underset{(0.095)}{0.053} \varepsilon_{t-1} - \underset{(0.095)}{0.399} \varepsilon_{t-2}$$

Commodity: Wool
Model: ARIMA(0,1,2) with constant
SBC = -40.754 , **Ljung-Box Q(12):** 3.771 (0.957)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.008)}{0.014} + v_t$$

$$v_t = \underset{(0.186)}{\varepsilon_t} - \underset{(0.094)}{0.172} \varepsilon_{t-1} - \underset{(0.094)}{0.420} \varepsilon_{t-2}$$

Commodity: Tobacco
Model: ARIMA(0,1,0) with constant
SBC = -100.611 , **Ljung-Box Q(12):**12.338 (0.419)
Degrees of freedom: 96

$$\Delta p_t = \underset{(0.014)}{0.003} + v_t$$

$$v_t = \underset{(0.141)}{\varepsilon_t}$$

Commodity: Rubber
Model: ARIMA(0,1,0) with constant
SBC = 37.095 , **Ljung-Box Q(12):**12.619 (0.397)
Degrees of freedom: 97

$$\Delta p_t = -\underset{(0.029)}{0.030} + v_t$$
$$v_t = \underset{(0.287)}{\varepsilon_t}$$

Commodity: Timber
Model: ARIMA(0,1,0) with constant
SBC = -75.495 , **Ljung-Box Q(12):**10.397 (0.581)
Degrees of freedom: 97

$$\Delta p_t = \underset{(0.016)}{0.008} + v_t$$
$$v_t = \underset{(0.162)}{\varepsilon_t}$$

Commodity: Copper
Model: ARIMA(0,1,0) with constant
SBC = -44.589 , **Ljung-Box Q(12):**8.658 (0.732)
Degrees of freedom: 97

$$\Delta p_t = -\underset{(0.019)}{0.009} + v_t$$
$$v_t = \underset{(0.189)}{\varepsilon_t}$$

Commodity: Aluminium
Model: ARIMA(1,1,2) with constant
SBC = -74.292 , **Ljung-Box Q(12):**2.965 (0.966)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.002)}{0.019} + v_t$$
$$v_t - \underset{(0.103)}{0.679} = \underset{(0.152)}{\varepsilon_t} - \underset{(n.a.)}{0.537\varepsilon_{t-1}} - \underset{(n.a.)}{0.463\varepsilon_{t-2}}$$

Commodity: Tin
Model: ARIMA(0,1,0) with constant
SBC = -43.686 , **Ljung-Box Q(12):**5.395 (0.943)
Degrees of freedom: 97

$$\Delta p_t = -\underset{(0.019)}{0.003} + v_t$$
$$v_t = \underset{(0.190)}{\varepsilon_t}$$

Commodity: Silver
Model: ARIMA(2,1,0) with constant
SBC = -46.430 , **Ljung-Box Q(12):**8.089 (0.620)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.014)}{0.003} + v_t$$
$$v_t - \underset{(0.098)}{0.041} v_{t-1} + \underset{(0.098)}{0.307} v_{t-2} = \underset{(0.181)}{\varepsilon_t}$$

Commodity: Lead
Model: ARIMA(0,1,0) with constant
SBC = -46.299 , **Ljung-Box Q(12):**13.818 (0.312)
Degrees of freedom: 97

$$\Delta p_t = -\underset{(0.019)}{0.008} + v_t$$
$$v_t = \underset{(0.188)}{\varepsilon_t}$$

Commodity: Zinc
Model: ARIMA(1,1,2) with constant
SBC = -22.495 , **Ljung-Box Q(12):**6.374 (0.702)
Degrees of freedom: 94

$$\Delta p_t = \underset{(0.002)}{0.000} + v_t$$
$$v_t - \underset{(0.138)}{0.478} v_{t-1} = \underset{(0.197)}{\varepsilon_t} - \underset{(n.a.)}{0.639} \varepsilon_{t-1} - \underset{(n.a.)}{0.361} \varepsilon_{t-2}$$

Appendix III.iv. Estimation Results for Relative Primary Product Price Series in Levels -minimum SBC specifications including dummies

This appendix gives details of the estimation results for relative primary product price series when outliers and structural breaks are taken into account. The price series are natural logarithms of the primary commodity price series shown relative to the Manufacturing Unit Value (MUV) index. Dummies were included for years with residuals beyond ± three standard deviations allowing for level shifts as well as single additive outliers. Price series were modelled for a constant and a linear trend allowing for a general ARMA specification for the error term, *i.e.*

$$p_t = a + \beta t + +\omega_{11} D_{1,19xx} + \omega_{21} D_{2,19xx} + ... + \omega_{1m} D_{m,19xx} + \omega_{2m} D_{2,19xx} + u_t$$

and:

$$u_t - \phi_1 u_{t-1} - \ldots - \phi_p u_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

where p_t is the relevant price series, α the coefficient on the constant, β the coefficient on the linear trend term, ϕ_i the coefficient on the relevant autoregressive error term (u_{t-i}), θ_i the coefficient on the i[th] moving average term (ε_{t-i}). ω_{1m} is the coefficient on the m[th] single outlier, with $D_{1,19xx}$ taking a value of one in the relevant year 19xx and zero otherwise. ω_{2m} is the coefficient on the level shift dummy $D_{2,19xx}$ marking a structural break in the relevant year 19xx. The subscript t indicates time period t and t-i the i[th] lag. ARMA models were estimated for all ARIMA(p,0,q) specifications such that p+q≤5 and the preferred model was identified by SBC. The estimation results were obtained using the ARIMA.SRC procedure in GAUSS, and are listed below. Ljung Box Q statistics for 12 autocorrelations are reported with P-values in parentheses. (As for the models without dummies possible autocorrelation is indicated for Cotton.)

Commodity: Coffee
Model: ARIMA(1,0,0), with constant and trend
SBC = 14.027 , **Ljung-Box Q(12):**7.296 (0.775)
Degrees of freedom: 94

$$p_t = -\underset{(0.399)}{0.792} - \underset{(0.007)}{0.006}t - \underset{(0.238)}{0.272}\, D_{1,1976} + \underset{(0.326)}{0.916} D_{2.1976} + u_t$$
$$u_t - \underset{(0.045)}{0.900} u_{t-1} = \underset{(0.235)}{\varepsilon_t}$$

Commodity: Cocoa
Model: ARIMA(1,0,1), with constant and trend
SBC = 6.076 , **Ljung-Box Q(12):**9.383 (0.587)
Degrees of freedom: 93

$$p_t = -\underset{(0.162)}{0.648} - \underset{(0.005)}{0.022}t - \underset{(0.203)}{0.456}\, D_{1,1947} + \underset{(0.261)}{1.403}\, D_{2.,1947} + u_t$$
$$u_t - \underset{(0.101)}{0.621} u_{t-1} = \underset{(0.223)}{\varepsilon_t} + \underset{(0.123)}{0.347}\varepsilon_{t-1}$$

Commodity: Tea
Model: ARIMA(1,0,1), with constant and trend
SBC = -79.279 , **Ljung-Box Q(12):**9.507 (0.485)
Degrees of freedom: 92

$$p_t = \underset{(0.121)}{0.153} - \underset{(0.002)}{0.000}t + \underset{(0.130)}{0.152}\, D_{1,1985} - \underset{(0.171)}{0.844}\, D_{2,1985} + u_t$$
$$u_t - \underset{(0.091)}{0.682} u_{t-1} = \underset{(0.144)}{\varepsilon_t} + \underset{(0.120)}{0.336}\varepsilon_{t-1}$$

Commodity: Rice
Model: ARIMA(1,0,1), with constant and trend
SBC = -59.409 , **Ljung-Box Q(12):**8.029 (0.626)
Degrees of freedom: 93

$$p_t = \underset{(0.301)}{0.915} - \underset{(0.006)}{0.021}t - \underset{(0.160)}{0.391}\, D_{1,1973} + \underset{(0.261)}{0.900}\, D_{2,1973} + u_t$$
$$u_t - \underset{(0.054)}{0.870} u_{t-1} = \underset{(0.159)}{\varepsilon_t} + \underset{(0.101)}{0.393}\varepsilon_{t-1}$$

Commodity: Wheat
Model: ARIMA(0,0,3), with constant and trend
SBC = -66.547 , **Ljung-Box Q(12):**5.396 (0.798)
Degrees of freedom: 92

$$p_t = \underset{(0.098)}{0.767} - \underset{(0.002)}{0.010}t + \underset{(0.111)}{0.167}\, D_{1,1973} - \underset{(0.145)}{0.090}\, D_{2,1973} + u_t$$
$$u_t = \underset{(0.151)}{\varepsilon_t} + \underset{(0.098)}{1.024}\, \varepsilon_{t-1} + \underset{(0.137)}{0.552}\, \varepsilon_{t-2} + \underset{(0.101)}{0.375}\, \varepsilon_{t-3}$$

Commodity: Maize
Model: ARIMA(1,0,0) with constant and trend
SBC = -17.486 , **Ljung-Box Q(12):**15.963 (0.143)
Degrees of freedom: 94

$$p_t = \underset{(0.153)}{0.787} - \underset{(0.003)}{0.011}t - \underset{(0.192)}{0.497}\, D_{1,1921} + \underset{(0.211)}{0.027}\, D_{2,1921} + u_t$$
$$u_t - \underset{(0.073)}{0.732}\, u_{t-1} = \underset{(0.202)}{\varepsilon_t}$$

Commodity: Sugar
Model: ARIMA(1,0,0) with constant and trend
SBC = 59.824 , **Ljung-Box Q(12):**9.638 (0.563)
Degrees of freedom: 90

$$p_t = \underset{(0.183)}{0.867} - \underset{(0.005)}{0.016}\, t + \underset{(0.269)}{0.597}\, D_{1,1963} + \underset{(0.270)}{0.305}\, D_{2,1963} + \underset{(0.279)}{0.670}\, D_{1,1974}$$
$$+ \underset{(0.307)}{0.219}\, D_{2,1974} + \underset{(0.271)}{0.908}\, D_{1,1980} - \underset{(0.293)}{0.249}\, D_{2,1980} + u_t$$
$$u_t - \underset{(0.084)}{0.650}\, u_{t-1} = \underset{(0.278)}{\varepsilon_t}$$

Commodity: Beef
Model: ARIMA(2,0,1), with constant and trend
SBC = -58.291 , **Ljung-Box Q(12):**5.036 (0.831)
Degrees of freedom: 88

$$p_t = -\underset{(0.081)}{1.435} - \underset{(0.005)}{0.019}\, t + \underset{(0.129)}{0.200}\, D_{1,1915} + \underset{(0.127)}{0.461}\, D_{2,1915} + \underset{(0.133)}{0.094}\, D_{1,1931}$$
$$+ \underset{(0.144)}{0.611}\, D_{2,1931} - \underset{(0.146)}{0.593}\, D_{1,1959} + \underset{(0.182)}{1.729}\, D_{2,1959} + u_t$$
$$u_t - \underset{(0.078)}{1.752}\, u_{t-1} + \underset{(0.071)}{0.837}\, u_{t-2} = \underset{(0.145)}{\varepsilon_t} - \underset{(n.a.)}{1.000}\, \varepsilon_{t-1}$$

Commodity: Lamb
Model: ARIMA(1,0,4), with constant and trend
SBC = -26.942 , **Ljung-Box Q(12):**4.151 (0.762)
Degrees of freedom: 86

$$p_t = -\underset{(0.326)}{2.124} + \underset{(0.007)}{0.011}\, t - \underset{(0.145)}{0.008}\, D_{1,1915} + \underset{(0.222)}{0.785}\, D_{2,1915} + \underset{(0.146)}{0.227}\, D_{1,1931}$$
$$+ \underset{(0.229)}{0.244}\, D_{2,1931} - \underset{(0.154)}{0.239}\, D_{1,1950} - \underset{(0.234)}{0.260}\, D_{2,1950} + u_t$$
$$u_t - \underset{(0.068)}{0.847}\, u_{t-1} = \underset{(0.165)}{\varepsilon_t} + \underset{(0.104)}{0.185}\, \varepsilon_{t-1} - \underset{(0.109)}{0.008}\, \varepsilon_{t-2} + \underset{(0.108)}{0.008}\, \varepsilon_{t-3} + \underset{(0.108)}{0.525}\, \varepsilon_{t-4}$$

Commodity: Bananas
Model: ARIMA(1,0,0) with constant and trend
SBC = -181.839 , **Ljung-Box Q(12):**8.438 (0.674)
Degrees of freedom: 96

$$p_t = \underset{(0.189)}{0.207} - \underset{(0.003)}{0.001}\, t + u_t$$
$$u_t - \underset{(0.037)}{0.926}\, u_{t-1} = \underset{(0.091)}{\varepsilon_t}$$

Commodity: Palm Oil
Model: ARIMA(1,0,1) with constant and trend
SBC = -25.496 , **Ljung-Box Q(12):**5.698 (0.840)
Degrees of freedom: 93

$$p_t = \underset{(0.097)}{0.377} - \underset{(0.002)}{0.006}\, t - \underset{(0.167)}{0.047}\, D_{1,1986} - \underset{(0.161)}{0.698}\, D_{2,1986} + u_t$$
$$u_t - \underset{(0.138)}{0.414}\, u_{t-1} = \underset{(0.190)}{\varepsilon_t} + \underset{(0.139)}{0.417}\, \varepsilon_{t-1}$$

Commodity: Cotton
Model: ARIMA(1,0,0) with constant and trend
SBC = -67.885 , **Ljung-Box Q(12):**20.823 (0.035)
Degrees of freedom: 96

$$p_t = \underset{(0.173)}{0.647} - \underset{(0.003)}{0.010}\, t + u_t$$
$$u_t - \underset{(0.057)}{0.833}\, u_{t-1} = \underset{(0.162)}{\varepsilon_t}$$

Commodity: Jute
Model: ARIMA(1,0,0) with constant and trend
SBC = -20.694 , **Ljung-Box Q(12):**15.672 (0.154)
Degrees of freedom: 94

$$p_t = \underset{(0.161)}{0.230} + \underset{(0.003)}{0.000}\, t - \underset{(0.191)}{0.032}\, D_{1,1986} - \underset{(0.216)}{0.949}\, D_{2,1986} + u_t$$
$$u_t - \underset{(0.067)}{0.760}\, u_{t-1} = \underset{(0.198)}{\varepsilon_t}$$

Commodity: Wool
Model: ARIMA(1,0,0) with constant and trend
SBC = -41.624 , **Ljung-Box Q(12)**:5.916 (0.879)
Degrees of freedom: 94

$$p_t = \underset{(0.181)}{1.173} - \underset{(0.004)}{0.013}\, t + \underset{(0.175)}{0.697}\, D_{1,1973} - \underset{(0.219)}{0.241}\, D_{2,1973} + u_t$$
$$u_t - \underset{(0.063)}{0.807}\, u_{t-1} = \underset{(0.178)}{\varepsilon_t}$$

Commodity: Tobacco
Model: ARIMA(3,0,2) with constant and trend
SBC = -115.239 , **Ljung-Box Q(12)**: 4.208 (0.755)
Degrees of freedom: 85

$$p_t = -\underset{(0.341)}{0.727} - \underset{(0.006)}{0.003}\, t + \underset{(0.104)}{0.053}\, D_{1,1920} + \underset{(0.150)}{0.314}\, D_{2,1920} + \underset{(0.102)}{0.093}\, D_{1,1960}$$
$$+ \underset{(0.145)}{0.572}\, D_{2,1960} - \underset{(0.107)}{0.003}\, D_{1,1996} - \underset{(0.160)}{0.363}\, D_{2,1960} + u_t$$
$$u_t - \underset{(0.040)}{2.373}\, u_{t-1} + \underset{(0.052)}{2.356}\, u_{t-2} - \underset{(0.034)}{0.959}\, u_{t-3} = \underset{(0.102)}{\varepsilon_t} - \underset{(n.a.)}{1.421}\, \varepsilon_{t-1} + \underset{(n.a.)}{\varepsilon_{t-2}}$$

Commodity: Rubber
Model: ARIMA(2,0,1) with constant and trend
SBC = 9.402 , **Ljung-Box Q(12)**:9.721 (0.374)
Degrees of freedom: 88

$$p_t = \underset{(0.057)}{2.439} - \underset{(0.003)}{0.035}\, t + \underset{(0.208)}{0.214}\, D_{1,1921} - \underset{(0.242)}{1.374}\, D_{2,1921} + \underset{(0.192)}{0.214}\, D_{1,1925}$$
$$+ \underset{(0.215)}{0.754}\, D_{2,1925} + \underset{(0.179)}{0.004}\, D_{1,1950} + \underset{(0.127)}{0.769}\, D_{2,1950} + u_t$$
$$u_t - \underset{(0.064)}{1.674}\, u_{t-1} + \underset{(0.060)}{0.791}\, u_{t-2} = \underset{(0.204)}{\varepsilon_t} - \underset{(n.a.)}{\varepsilon_{t-1}}$$

Commodity: Timber
Model: ARIMA(1,0,0) with constant and trend
SBC = -88.961 , **Ljung-Box Q(12)**:10.407 (0.494)
Degrees of freedom: 94

$$p_t = -\underset{(0.098)}{0.993} + \underset{(0.002)}{0.011}\, t + \underset{(0.139)}{0.434}\, D_{1,1993} + \underset{(0.157)}{0.036}\, D_{2,1993} + u_t$$
$$u_t - \underset{(0.075)}{0.720}\, u_{t-1} = \underset{(0.141)}{\varepsilon_t}$$

Commodity: Copper
Model: ARIMA(1,0,0) with constant and trend
SBC = -42.081 , **Ljung-Box Q(12):**6.542 (0.835)
Degrees of freedom: 96

$$p_t = \underset{(0.225)}{0.243} - \underset{(0.004)}{0.004}\, t + u_t$$
$$u_t - \underset{(0.052)}{0.856}\, u_{t-1} = \underset{(0.184)}{\varepsilon_t}$$

Commodity: Aluminium
Model: ARIMA(1,0,1) with constant and trend
SBC = -76.272 , **Ljung-Box Q(12):**6.497 (0.772)
Degrees of freedom: 93

$$p_t = \underset{(0.285)}{1.050} - \underset{(0.005)}{0.026}\, t - \underset{(0.148)}{0.303}\, D_{1,1915} + \underset{(0.239)}{0.848}\, D_{2,1915} + u_t$$
$$u_t - \underset{(0.054)}{0.874}\, u_{t-1} = \underset{(0.146)}{\varepsilon_t} + \underset{(0.105)}{0.367}\, \varepsilon_{t-1}$$

Commodity: Tin
Model: ARIMA(1,0,0) with constant and trend
SBC = -57.994 , **Ljung-Box Q(12):**8.810 (0.639)
Degrees of freedom: 94

$$p_t = -\underset{(0.125)}{1.194} + \underset{(0.002)}{0.010}\, t + \underset{(0.158)}{0.123}\, D_{1,1986} - \underset{(0.173)}{1.009}\, D_{2,1986} + u_t$$
$$u_t - \underset{(0.069)}{0.741} = \underset{(0.164)}{\varepsilon_t}$$

Commodity: Silver
Model: ARIMA(1,0,0) with constant and trend
SBC = -46.721 , **Ljung-Box Q(12):**13.712 (0.249)
Degrees of freedom: 92

$$p_t = -\underset{(0.265)}{0.744} + \underset{(0.006)}{0.003}\, t + \underset{(0.176)}{0.659}\, D_{1,1979} - \underset{(0.252)}{0.099}\, D_{2,1979} + \underset{(0.169)}{0.108}\, D_{1,1980}$$
$$- \underset{(0.232)}{0.147}\, D_{2,1980} + u_t$$
$$u_t - \underset{(0.052)}{0.891}\, u_{t-1} = \underset{(0.167)}{\varepsilon_t}$$

Commodity: Lead
Model: ARIMA(2,0,0) with constant and trend
SBC = 51.996 , **Ljung-Box Q(12):**4.087 (0.943)
Degrees of freedom: 93

$$p_t = -\underset{(0.090)}{0.187} - \underset{(0.002)}{0.001}\, t + \underset{(0.144)}{0.699}\, D_{1,1979} - \underset{(0.136)}{0.517}\, D_{2,1979} + u_t$$
$$u_t - \underset{(0.102)}{0.827}\, u_{t-1} + \underset{(0.102)}{231}\, u_{t-2} = \underset{(0.166)}{\varepsilon_t}$$

Commodity: Zinc
Model: ARIMA(1,0,0) with constant and trend
SBC = -13.208 , **Ljung-Box Q(12):**11.972 (0.366)
Degrees of freedom: 92

$$p_t = \underset{(0.127)}{0.069} - \underset{(0.004)}{0.004}\, t + \underset{(0.192)}{0.128}\, D_{1,1915} + \underset{(0.197)}{0.322}\, D_{2,1915} + \underset{(0.193)}{0.026}\, D_{1,1973}$$

$$+ \underset{(0.198)}{0.088}\, D_{2,1973} + u_t$$

$$u_t - \underset{(0.078)}{0.664}\, u_{t-1} = \underset{(0.199)}{\varepsilon_t}$$

Appendix III.v. Estimation Results for Relative Primary Product Price Series in First Differences -minimum SBC specifications including dummies.

This appendix gives details of the estimation results for relative primary product price series. The price series are natural logarithms of the primary commodity price series shown relative to the Manufacturing Unit Value (MUV) index. Dummies were included for years with residuals beyond ± three standard deviations allowing for level shifts as well as single additive outliers. Price series were modelled including a constant and allowing for a general ARMA specification for the error term, *i.e.*

$$\Delta p_t = \beta + \omega_{11}\Delta D_{1,19xx} + \omega_{21}\Delta D_{2,19xx} + ... + \omega_{1m}\Delta D_{m,19xx} + \omega_{2m}\Delta D_{2,19xx} + v_t$$

$$v_t - \phi_1 v_{t-1} - ... - \phi_p v_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - ... - \theta_q \varepsilon_{t-q}$$

where Δp_t is the relevant price series in first differences, β the coefficient on the constant (*i.e.* the drift term), ϕ_i the coefficient on the relevant autoregressive error term (v_{t-i}), θ_i the coefficient on the ith moving average term (ε_{t-i}). ω_{1m} is the coefficient on the first difference of the mth single additive outlier. ω_{2m} is the coefficient on the first difference in the level shift dummy $D_{2,19xx}$ marking a structural break in the relevant year 19xx. The subscript t indicates time period t and t-i the ith lag. ARIMA models were estimated for all ARIMA(p,1,q) specifications such that p+q≤5 and the preferred model was identified by SBC. The estimation results were obtained using the ARIMA.SRC procedure in GAUSS, and are listed below. Ljung Box Q statistics for 12 autocorrelations are reported with P-values in parentheses. (At the 5% significance level this now indicates possible autocorrelation for Wheat and Jute as well as Tobacco.)

Commodity: Coffee
Model: ARIMA(0,1,0) with constant
SBC = 7.772 , **Ljung-Box Q(12):**8.687 (0.729)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.024)}{0.010} - \underset{(0.240)}{0.412}\,\Delta D_{1,1976} + \underset{(0.341)}{1.190}\,\Delta D_{2,1976} + v_t$$
$$v_t = \underset{(0.238)}{\varepsilon_t}$$

Commodity: Cocoa
Model: ARIMA(2,1,0) with constant
SBC = 7.622 , **Ljung-Box Q(12):**7.704 (0.658)
Degrees of freedom: 93

$$\Delta p_t = -\underset{(0.019)}{0.021} - \underset{(0.220)}{0.385}\,\Delta D_{1,1947} + \underset{(0.325)}{1.170}\,\Delta D_{2,1947} + v_t$$
$$v_t - \underset{(0.098)}{0.058}\,v_{t-1} + \underset{(0.098)}{0.330}\,v_{t-2} = \underset{(0.229)}{\varepsilon_t}$$

Commodity: Tea
Model: ARIMA(0,1,0) with constant
SBC = -76.904 , **Ljung-Box Q(12):**16.098 (0.187)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.016)}{0.002} + \underset{(0.155)}{0.191}\,\Delta D_{1,1985} - \underset{(0.220)}{0.754}\,\Delta D_{2,1985} + v_t$$
$$v_t = \underset{(0.154)}{\varepsilon_t}$$

Commodity: Rice
Model: ARIMA(1,1,1) with constant
SBC = -64.595 , **Ljung-Box Q(12):**6.870 (0.738)
Degrees of freedom: 93

$$\Delta p_t = -\underset{(0.019)}{0.021} - \underset{(0.154)}{0.485}\,\Delta D_{1,1973} + \underset{(0.261)}{1.117}\,\Delta D_{2,1973} + v_t$$
$$v_t + \underset{(0.163)}{0.547}\,v_{t-1} = \underset{(0.159)}{\varepsilon_t} + \underset{(0.108)}{0.841}\,\varepsilon_{t-1}$$

Commodity: Wheat
Model: ARIMA(0,1,2) with constant
SBC = -69.333 , **Ljung-Box Q(12)**:22.965 (0.011)
Degrees of freedom: 93

$$\Delta p_t = -\underset{(0.011)}{0.015} - \underset{(0.137)}{0.108}\,\Delta D_{1,1973} + \underset{(0.223)}{0.492}\,\Delta D_{2,1973} + v_t$$
$$v_t = \underset{(0.155)}{\varepsilon_t} + \underset{(0.092)}{0.153}\,\varepsilon_{t-1} - \underset{(0.092)}{0.479}\,\varepsilon_{t-2}$$

Commodity: Maize
Model: ARIMA(0,1,2) with constant
SBC = -23.026 , **Ljung-Box Q(12)**:15.672 (0.109)
Degrees of freedom: 93

$$\Delta p_t = -\underset{(0.009)}{0.005} - \underset{(0.186)}{0.264}\,\Delta D_{1,1921} - \underset{(0.215)}{0.365}\,\Delta D_{2,1921} + v_t$$
$$v_t = \underset{(0.196)}{\varepsilon_t} - \underset{(0.099)}{0.237}\,\varepsilon_{t-1} - \underset{(0.100)}{0.356}\,\varepsilon_{t-2}$$

Commodity: Sugar
Model: ARIMA(0,1,2) with constant
SBC = 59.473 , **Ljung-Box Q(12)**:8.033 (0.626)
Degrees of freedom: 89

$$\Delta p_t = -\underset{(0.017)}{0.010} - \underset{(0.264)}{0.205}\,\Delta D_{1,1921} - \underset{(0.386)}{0.744}\,\Delta D_{2,1921} + \underset{(0.252)}{0.43}\ 6\Delta D_{1,1963}$$
$$+\underset{(0.364)}{0.220}\,\Delta D_{2,1963} + \underset{(0.258)}{0.585}\,\Delta D_{1,1974} + \underset{(0.383)}{0.324}\,\Delta D_{2,1974} + v_t$$
$$v_t = \underset{(0.278)}{\varepsilon_t} + \underset{(0.102)}{0.005}\,\varepsilon_{t-1} - \underset{(0.103)}{0.447}\,\varepsilon_{t-2}$$

Commodity: Beef
Model: ARIMA(0,1,0) with constant
SBC = -57.689 , **Ljung-Box Q(12)**:10.451 (0.576)
Degrees of freedom: 91

$$\Delta p_t = -\underset{(0.017)}{0.016} + \underset{(0.160)}{0.111}\,\Delta D_{1,1915} + \underset{(0.227)}{0.638}\,\Delta D_{2,1915} + \underset{(0.160)}{0.116}\,\Delta D_{1,1931}$$
$$+\underset{(0.227)}{0.609}\,\Delta D_{2,1931} - \underset{(0.160)}{0.250}\,\Delta D_{1,1959} + \underset{(0.227)}{1.091}\,\Delta D_{2,1959} + v_t$$
$$v_t = \underset{(0.159)}{\varepsilon_t}$$

Commodity: Lamb
Model: ARIMA(0,1,0) with constant
SBC = -31.749 , **Ljung-Box Q(12):**13.031 (0.367)
Degrees of freedom: 93

$$\Delta p_t = \underset{(0.019)}{0.003} + \underset{(0.189)}{0.178}\,\Delta D_{1,1915} + \underset{(0.269)}{0.624}\,\Delta D_{2,1915} + \underset{(0.189)}{0.081}\,\Delta D_{1,1931}$$
$$+ \underset{(0.269)}{0.633}\,\Delta D_{2,1931} + v_t$$
$$v_t = \underset{(0.188)}{\varepsilon_t}$$

Commodity: Bananas
Model: ARIMA(0,1,0) with constant
SBC = -187.437 , **Ljung-Box Q(12):**10.778 (0.548)
Degrees of freedom: 97

$$\Delta p_t = \underset{(0.009)}{0.000} + v_t$$
$$v_t = \underset{(0.091)}{\varepsilon_t}$$

Commodity: Palm Oil
Model: ARIMA(1,1,2) with constant
SBC = -20.492 , **Ljung-Box Q(12):**6.644 (0.674)
Degrees of freedom: 92

$$\Delta p_t = -\underset{(0.002)}{0.005} - \underset{(0.199)}{0.008}\,\Delta D_{1,1986} - \underset{(0.291)}{0.782}\,\Delta D_{2,1986} + v_t$$
$$v_t - \underset{(0.147)}{0.510}\,v_{t-1} = \underset{(0.192)}{\varepsilon_t} - \underset{(n.a.)}{0.625}\,\varepsilon_{t-1} - \underset{(n.a.)}{0.375}\,\varepsilon_{t-2}$$

Commodity: Cotton
Model: ARIMA(2,1,2) with constant
SBC = -73.914 , **Ljung-Box Q(12):**9.176 (0.328)
Degrees of freedom: 93

$$\Delta p_t = -\underset{(0.010)}{0.008} + v_t$$
$$v_t - \underset{(0.077)}{1.303}\,v_{t-1} + \underset{(0.076)}{0.758}\,v_{t-2} = \underset{(0.149)}{\varepsilon_t} - \underset{(0.061)}{1.649}\,\varepsilon_{t-1} + \underset{(0.065)}{0.964}\,\varepsilon_{t-2}$$

Commodity: Jute
Model: ARIMA(0,1,0) with constant
SBC = -17.824 , **Ljung-Box Q(12):**21.972 (0.038)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.021)}{0.000} - \underset{(0.210)}{0.085}\,\Delta D_{1,1986} - \underset{(0.299)}{0.849}\,\Delta D_{2,1986} + v_t$$
$$v_t = \underset{(0.209)}{\varepsilon_t}$$

Commodity: Wool
Model: ARIMA(0,1,0) with constant
SBC = -43.814 , **Ljung-Box Q(12):**10.315 (0.588)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.019)}{0.015} + \underset{(0.184)}{0.518}\,\Delta D_{1,1973} + \underset{(0.262)}{0.110}\,\Delta D_{2,1973} + v_t$$
$$v_t = \underset{(0.183)}{\varepsilon_t}$$

Commodity: Tobacco
Model: ARIMA(0,1,0) with constant
SBC = -120.842 , **Ljung-Box Q(12):**28.186 (0.005)
Degrees of freedom: 92

$$\Delta p_t = \underset{(0.012)}{0.000} + \underset{(0.119)}{0.122}\,\Delta D_{1,1960} + \underset{(0.169)}{0.568}\,\Delta D_{2,1960}$$
$$-\underset{(0.119)}{0.054}\,\Delta D_{1,1996} - \underset{(0.169)}{0.315}\,\Delta D_{2,1996} + v_t$$
$$v_t = \underset{(0.118)}{\varepsilon_t}$$

Commodity: Rubber
Model: ARIMA(0,1,0) with constant
SBC = 23.822 , **Ljung-Box Q(12):**15.660 (0.207)
Degrees of freedom: 93

$$\Delta p_t = -\underset{(0.026)}{0.051} + \underset{(0.251)}{0.281}\,\Delta D_{1,1925} + \underset{(0.357)}{0.776}\,\Delta D_{2,1925} - \underset{(0.251)}{0.244}\,\Delta D_{1,1950}$$
$$+\underset{(0.357)}{1.239}\,\Delta D_{2,1950} + v_t$$
$$v_t = \underset{(0.249)}{\varepsilon_t}$$

Commodity: Timber
Model: ARIMA(0,1,0) with constant
SBC = -85.915 , **Ljung-Box Q(12):**11.695 (0.470)
Degrees of freedom: 95

$$\Delta p_t = \underset{(0.015)}{0.005} + \underset{(0.149)}{0.277}\,\Delta D_{1,1993} + \underset{(0.211)}{0.342}\,\Delta D_{2,1993} + v_t$$
$$v_t = \underset{(0.148)}{\varepsilon_t}$$

Commodity: Copper
Model: ARIMA(0,1,0) with constant
SBC = -46.908 , **Ljung-Box Q(12):**14.206 (0.288)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.018)}{0.004} - \underset{(0.181)}{0.115}\,\Delta D_{1,1975} - \underset{(0.258)}{0.496}\,\Delta D_{2,1975} + v_t$$
$$v_t = \underset{(0.180)}{\varepsilon_t}$$

Commodity: Aluminium
Model: ARIMA(0,1,1) with constant
SBC = -81.302 , **Ljung-Box Q(12):**6.385 (0.846)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.020)}{0.030} - \underset{(0.149)}{0.437}\,\Delta D_{1,1915} + \underset{(0.243)}{1.101}\,\Delta D_{2,1915} + v_t$$
$$v_t = \underset{(0.149)}{\varepsilon_t} + \underset{(0.100)}{0.309}\,\varepsilon_{t-1}$$

Commodity: Tin
Model: ARIMA(0,1,0) with constant
SBC = -53.899 , **Ljung-Box Q(12):**9.804 (0.633)
Degrees of freedom: 95

$$\Delta p_t = \underset{(0.018)}{0.005} + \underset{(0.175)}{0.020}\,\Delta D_{1,1986} - \underset{(0.249)}{0.817}\,\Delta D_{2,1986} + v_t$$
$$v_t = \underset{(0.174)}{\varepsilon_t}$$

Commodity: Silver
Model: ARIMA(0,1,0) with constant
SBC = -57.199 , **Ljung-Box Q(12):**18.105 (0.113)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.017)}{0.015} - \underset{(0.172)}{0.543}\,\Delta D_{1,1979} + \underset{(0.245)}{1.153}\,\Delta D_{2,1979} + v_t$$
$$v_t = \underset{(0.171)}{\varepsilon_t}$$

Commodity: Lead
Model: ARIMA(0,1,0) with constant
SBC = -46.299 , **Ljung-Box Q(12):**13.818 (0.312)
Degrees of freedom: 97

$$\Delta p_t = -\underset{(0.019)}{0.008} + v_t$$
$$v_t = \underset{(0.188)}{\varepsilon_t}$$

Commodity: Zinc
Model: ARIMA(1,1,2) with constant
SBC = -10.568 , **Ljung-Box Q(12):**5.923 (0.748)
Degrees of freedom: 90

$$\Delta p_t = -\underset{(0.004)}{0.008} + \underset{(0.169)}{0.005}\,\Delta D_{1,1915} + \underset{(0.175)}{0.356}\,\Delta D_{2,1915}$$

$$-\underset{(0.181)}{0.089}\,\Delta D_{1,1973} + \underset{(0.216)}{0.384}\,\Delta D_{2,1973} + v_t$$

$$v_t - \underset{(0.154)}{0.538}\,v_{t-1} = \underset{(0.195)}{\varepsilon_t} - \underset{n.a.}{0.709}\,\varepsilon_{t-1} - \underset{(n.a.)}{0.291}\,\varepsilon_{t-2}$$

Appendix III.vi Normality Tests for Residuals from ARIMA Estimates of Relative Primary Commodity Prices with and without Dummy Variables.

To assess whether the residuals obtained from the ARIMA estimates reported in the main text are normally distributed, Bowman Shelton normality tests were conducted for the residuals from the selected ARIMA models in levels and first differences, before and after outliers have been accounted for. Following Newbold (1995), the Bowman- Shelton test Statistic is defined as follows:

$$B = n\left[\frac{(Skewness)^2}{6} + \frac{(Kurtosis-3)^2}{24}\right]$$

and Skewness is defined as:

$$\text{Skewness} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^3/n}{S^3}$$

and for Kurtosis:

$$\text{Kurtosis} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^4/n}{S^4}$$

where B is the Bowman-Shelton test statistic, n the number of observations, x_i the i^{th} observation for variable x and $\bar{x}$ is the average value of variable x.

The test results obtained for ARIMA models without dummies are given in table III.vi.i below.

Table III.vi.i. Bowman-Shelton Test Statistics for (no dummies)

Commodity	Normality test for model in levels	Normality test for model in first differences
Coffee	2.784	5.969!
Cocoa	8.039!	10.938!
Tea	9.988!	12.581!
Rice	13.991!	7.054!
Wheat	10.350!	12.272!
Maize	16.032!	15.560!
Sugar	91.814!	38.743!
Beef	69.857!	88.980!
Lamb	65.033!	53.623!
Banana	0.033	0.014
Palm Oil	3.559	17.563!
Cotton	0.033	0.268
Jute	15.076!	2.270
Wool	4.622!	4.929!
Tobacco	166.304!	125.040!
Rubber	21.056!	29.884!
Timber	26.292!	15.431!
Copper	0.458	1.651
Aluminium	45.254!	39.851!
Tin	31.543!	39.298!
Silver	44.813!	8.990!
Lead	0.350	0.044
Zinc	156.732!	100.771!

Values of the Bowman-Shelton test statistic indicating non-normality are indicated by !

Given a five percent critical value of 4.29 for the rejection for the null hypothesis of normality, the residuals of 17 of the 23 commodities covered in the estimates in levels appear to be distributed non-normally. For the estimates in first differences, a total of 18 residual series are shown to be non-normally distributed.

If outliers are accounted for, a larger number of ARIMA models seem to have normally distributed residuals. The Bowman-Shelton test results for residuals from ARIMA models with dummies are summarised below in table III.vi.ii.

Table III.vi.ii Bowman-Shelton Test Statistics (including dummies)

Commodity	Normality test for model in levels	Normality test for model in first differences
Coffee	0.295	3.820
Cocoa	0.752	1.594
Tea	1.969	14.47!
Rice	7.361!	4.271
Wheat	2.472	1.424
Maize	5.645!	6.771!
Sugar	9.158!	4.153
Beef	2.152	14.505!
Lamb	15.120!	29.916!
Banana	0.033	0.014
Palm Oil	6.115!	1.935
Cotton	0.033	0.268
Jute	2.611	2.251
Wool	0.171	2.808
Tobacco	2.368	7.502!
Rubber	0.175	3.867
Timber	2.981	1.391
Copper	0.458	1.982
Aluminium	9.801!	14.639!
Tin	3.682	4.620!
Silver	37.240!	17.565!
Lead	2.837	0.044
Zinc	107.433!	70.141!

Values of the Bowman-Shelton test statistic indicating non-normality are indicated by !

It can be seen from table III.vi.ii. that when outliers are accounted for, only eight of the 23 residual series are shown to be non-normally distributed for ARIMA estimates in levels. For residuals from estimates in first differences, the distributions appear to be non-normal at the 5 percent significance level in nine cases.

Appendix III.vii Confidence Intervals for Coefficient Estimates

This Appendix lists 90% and 95% confidence intervals for the coefficient estimates reported in the main text for the models with and without dummy variables. Table III.vii.i. provides the confidence intervals, assuming that the asymptotic properties of the estimator hold and the 90% confidence interval is comprised between ±1.65 standard deviations around the sample mean. As in the main Chapter, the values of the interval limits are multiplied by 100.

Table III.vii.i. 90% Confidence Intervals for Coefficient Estimates.

Commodity	**Trend** (Levels) (No Dummies)	**Drift** (First Differences) (No Dummies)	**Trend** (Levels) (Incl. Dummies)	**Drift** (First Differences) (Incl. Dummies)
Coffee	[-0.198, 1.060]	[-3.964, 4.436]	[-1.806, 0.653]	[-4.993, 3.036]
Cocoa	[-1.226, 0.608]	[-4.247, 2.404]	[-2.953, -1.425]	[-5.199, 0.936]
Tea†	[-1.398, -0.091]	[-3.699, 1.791]	[-0.431, 0.346]	[-2.785, 2.431]
Rice	[-1.460, -0.759]	[-1.985, -0.377]	[-3.043, -1.143]	[-5.306, 1.034]
Wheat	[-1.303, -0.798]	[-2.374, 0.420]	[-1.321, -0.579]	[-3.263, 0.327]
Maize	[-1.415, -0.614]	[-2.173, 0.225]	[-1.587, -0.520]	[-1.960, 0.864]
Sugar	[-1.496, -0.639]	[-3.589, 1.160]	[-2.354, -0.772]	[-3.908, 1.867]
Beef	[0.381, 2.331]	[-2.630, 4.243]	[-2.718, -0.995]	[-4.312, 1.153]
Lamb	[1.236, 2.418]	[-2.025, 5.121]	[-0.035, 2.322]	[-2.933, 3.464]
Banana	[-0.621, 0.406]	[-1.482, 1.562]	[-0.621, 0.406]	[-1.482, 1.562]
Palm Oil	[-1.430, -0.641]	[-3.455, 2.029]	[-0.875, -0.245]	[-0.845, -0.118]
Cotton	[-1.472, -0.496]	[-2.479, 0.908]	[-1.472, -0.496]	[-2.479, 0.908]
Jute	[-1.429, 0.037]	[-2.786, 1.186]	[-0.462, 0.553]	[-3.567, 3.478]
Wool	[-2.128, -1.013]	[-2.743, -0.157]	[-1.952, -0.679]	[-4.630, 1.541]
Tobacco†	[-0.878, 1.816]	[-2.097, 2.641]	[-1.327, 0.722]	[-2.016, 2.037]
Rubber	[-3.530, -2.145]	[-7.805, 1.765]	[-3.953, -3.012]	[-9.322, -0.829]
Timber	[0.881, 1.394]	[-1.890, 3.499]	[0.799, 1.386]	[-2.033, 2.945]
Copper	[-1.047, 0.211]	[-4.092, 2.217]	[-1.047, 0.211]	[-3.468, 2.606]
Aluminium	[-2.214, -1.528]	[-2.315, -1.517]	[-3.424, -1.791]	[-6.248, 0.268]
Tin	[-0.689, 0.882]	[-3.499, 2.839]	[0.575, 1.368]	[-2.427, 3.434]
Silver	[-0.717, 0.762]	[-2.718, 2.056]	[-0.659, 1.172]	[-4.373, 1.390]
Lead	[-1.054, -0.142]	[-3.911, 2.343]	[-0.403, 0.234]	[-3.911, 2.343]
Zinc	[-0.206, 0.324]	[-0.291, 0.364]	[-1.117, 0.299]	[-1.532, -0.165]

Aside from the somewhat narrower interval width -which is to be expected in this case- most confidence intervals on the trend or drift coefficients imply the same conclusions on statistical significance as were reported in the main text for t-tests with nominal 5% critical values. There are two exceptions to this, however. The trend coefficient for Tea, when modelling the series without accounting for structural breaks, is identified as negative with 90% confidence. Likewise, the drift coefficient for Wool, again modelling the price series without accounting for outliers or structural breaks, would be shown to be negative if inferences were based on 10% critical values or 90% confidence intervals, while the coefficient sign would be indeterminate for the 95% confidence interval reported in the main text.

One could argue that given the degrees of freedom observed for the regression models obtained through use of the SBC it would be more appropriate to use critical values of ±1.99 for two tailed tests with a 5% rejection region and ±1.66 for two tailed tests with a 10% rejection region. This would of course have repercussions on the exact limits of the confidence intervals obtained. The 95% confidence intervals for the estimated trend and drift coefficients would in this case be given by the intervals in the table below:

Table III.vii.ii. 95% Confidence Intervals for Coefficient Estimates (exact interval width).

Commodity	Trend (Levels) (No Dummies)	Drift (First Differences) (No Dummies)	Trend (Levels) (Incl. Dummies)	Drift (First Differences) (Incl. Dummies)
Coffee	[-0.328, 1.190]	[-4.829, 5.301]	[-2.060, 0.907]	[-5.820, 3.863]
Cocoa	[-1.415, 0.797]	[-4.933, 3.089]	[-3.111, -1.267]	[-5.831, 1.568]
Tea†	[-1.532, 0.044]	[-4.265, 2.357]	[-0.511, 0.426]	[-3.323, 2.968]
Rice	[-1.532, -0.687]	[-2.151, -0.211]	[-3.238, -0.947]	[-5.959, 1.688]
Wheat	[-1.355, -0.746]	[-2.662, 0.708]	[-1.398, -0.503]	[-3.633, 0.697]
Maize	[-1.498, -0.532]	[-2.420, 0.472]	[-1.697, -0.410]	[-2.251, 1.155]
Sugar	[-1.585, -0.550]	[-4.078, 1.649]	[-2.517, -0.609]	[-4.502, 2.462]
Beef	[0.180, 2.531]	[-3.338, 4.951]	[-2.896, -0.817]	[-4.875, 1.716]
Lamb	[1.114, 2.540]	[-2.761, 5.857]	[-0.278, 2.565]	[-3.592, 4.123]
Banana	[-0.727, 0.512]	[-1.796, 1.875]	[-0.727, 0.512]	[-1.796, 1.875]
Palm Oil	[-1.511, -0.560]	[-4.020, 2.594]	[-0.940, -0.180]	[-0.919, -0.043]
Cotton	[-1.572, -0.395]	[-2.828, 1.257]	[-1.572, -0.395]	[-2.828. 1.257]
Jute	[-1.580, 0.188]	[-3.195, 1.596]	[-0.567, 0.657]	[-4.293, 4.204]
Wool	[-2.242, -0.899]	[-3.009, 0.110]	[-2.084, -0.548]	[-5.266, 2.176]
Tobacco†	[-1.156, 2.094]	[-2.586, 3.130]	[-1.538, 0.933]	[-2.433, 2.455]
Rubber	[-3.672, -2.003]	[-8.791, 2.751]	[-4.050, -2.914]	[-10.197, 0.046]
Timber	[0.828, 1.447]	[-2.445, 4.054]	[0.738, 1.447]	[-2.546, 3.458]
Copper	[-1.177, 0.340]	[-4.742, 2.867]	[-1.177, 0.340]	[-4.094, 3.231]
Aluminium	[-2.285, -1.457]	[-2.397, -1.435]	[-3.592, -1.623]	[-6.919, 0.939]
Tin	[-0.851, 1.043]	[-4.152, 3.492]	[0.494, 1.449]	[-3.031, 4.038]
Silver	[-0.870, 0.914]	[-3.210, 2.548]	[-0.847, 1.361]	[-4.967, 1.984]
Lead	[-1.148, -0.049]	[-4.555, 2.987]	[-0.468, 0.299]	[-4.555, 2.987]
Zinc	[-0.260, 0.379]	[-0.358, 0.431]	[-1.263, 0.445]	[-1.673, -0.024]

The corresponding 90% confidence intervals would be:

Table III.vii.iii. 90% Confidence Intervals for Coefficient Estimates (exact interval width).

Commodity	Trend (Levels) (No Dummies)	Drift (First Differences) (No Dummies)	Trend (Levels) (Incl. Dummies)	Drift (First Differences) (Incl. Dummies)
Coffee	[-0.202, 1.064]	[-3.989, 4.461]	[-1.814, 0.661]	[-5.017, 3.060]
Cocoa	[-1.231, 0.614]	[-4.267, 2.424]	[-2.958, -1.420]	[-5.217, 0.955]
Tea†	[-1.401, -0.087]	[-3.716, 1.808]	[-0.434, 0.348]	[-2.801, 2.447]
Rice	[-1.462, -0.757]	[-1.990, -0.372]	[-3.048, -1.137]	[-5.325, 1.054]
Wheat	[-1.305, -0.797]	[-2.383, 0.428]	[-1.324, -0.577]	[-3.274, 0.338]
Maize	[-1.418, -0.612]	[-2.181, 0.232]	[-1.597, -0.517]	[-1.968, 0.872]
Sugar	[-1.499, -0.636]	[-3.603, 1.174]	[-2.359, -0.767]	[-3.925, 1.884]
Beef	[0.375, 2.337]	[-2.651, 4.264]	[-2.724, -0.989]	[-4.328, 1.169]
Lamb	[1.233, 2.422]	[-2.047, 5.142]	[-0.042, 2.329]	[-2.952, 3.483]
Banana	[-0.624, 0.409]	[-1.491, 1.571]	[-0.624, 0.409]	[-1.491, 1.571]
Palm Oil	[-1.432, -0.639]	[-3.472, 2.045]	[-0.877, -0.243]	[-0.847, -0.116]
Cotton	[-1.475, -0.493]	[-2.489, 0.918]	[-1.475, -0.493]	[-2.489, 0.918]
Jute	[-1.433, 0.042]	[-2.798, 1.198]	[-0.465, 0.556]	[-3.589, 3.500]
Wool	[-2.131, -1.010]	[-2.751, -0.149]	[-1.956, -0.675]	[-4.649, 1.559]
Tobacco†	[-0.887, 1.824]	[-2.112, 2.656]	[-1.333, 0.728]	[-2.028, 2.037]
Rubber	[-3.534, -2.141]	[-7.834, 1.794]	[-3.956, -3.009]	[-9.348, -0.803]
Timber	[0.879, 1.395]	[-1.906, 3.515]	[0.797, 1.388]	[-2.048, 2.960]
Copper	[-1.051, 0.214]	[-4.111, 2.236]	[-1.051, 0.214]	[-3.487, 2.624]
Aluminium	[-2.216, -1.526]	[-2.317, -1.515]	[-3.429, -1.787]	[-6.267, 0.287]
Tin	[-0.694, 0.886]	[-3.518, 2.858]	[0.573, 1.370]	[-2.445, 3.452]
Silver	[-0.722, 0.766]	[-2.732, 2.071]	[-0.664, 1.178]	[-4.391, 1.407]
Lead	[-1.057, -0.140]	[-3.930, 2.362]	[-0.405, 0.235]	[-3.930, 2.362]
Zinc	[-0.207, 0.326]	[-0.293, 0.366]	[-1.121, 0.304]	[-1.536, -0.161]

Again there are generally no differences aside from the predictably somewhat larger interval width. The only exception here is the confidence interval for the drift coefficient of Rubber prices when the price series is modelled as difference stationary including dummy variables for structural breaks and outliers. This confidence interval now includes zero implying that the coefficient estimate should be regarded as statistically insignificant if a conventional t-test is applied. It is worth noting though that the confidence interval was rather wide under the assumption that the asymptotic properties hold and that the upper limits of the 90% and 95% confidence intervals were close to zero in any case. Against this background, and considering the fact that this was the only case were conclusions on the significance of the coefficient estimate were affected, no problems are expected if asymptotic critical values are used for t-tests at this point and in subsequent chapters.

Chapter 4

Further Attempts at Assessing the Significance of Trend Coefficient Estimates

Chapter 4: Further Attempts at Assessing the Significance of Trend Coefficient Estimates

As pointed out in chapter 3, it is generally the case that the estimates for trend and drift coefficients[1] for the commodity price series reported above are reasonably similar, while the differences in the absolute values of the corresponding t-ratios are noticeably larger. In the present chapter, those series for which significant trend coefficients were obtained under the trend stationary specification while the corresponding estimate for the drift coefficient in the difference stationary model turned out to be insignificant are further investigated. In addition, different general approaches towards model specification and trend testing are undertaken for all the commodities covered. As a first step in this investigation the model specifications selected previously under the Schwarz-Bayesian criterion (SBC) were used to generate simulated data series to which the competing models in levels and first differences were fitted subsequently. This was done in all those cases where conclusions about the presence of a trend depended on assumptions about the order of integration of the series.

[1] It should be recalled that, where reference is made to a drift term in contrast to a trend coefficient, the term trend coefficient is used to refer to the estimated magnitude for a secular trend obtained from a trend stationary model, while the term drift coefficient is used in reference to the estimated magnitude of the average change of a series obtained from a difference stationary model.

4.1. Simulated Data Series for Selected Commodities

4.1.1. Simulation Methodology

The simulated series were compiled using a 32-bit pseudo random number generator to generate a series of random numbers with a standard normal distribution[2]. The generated random numbers were then placed in a rx1 column vector, where r is the number of rows, and were subsequently multiplied by the standard error of the residuals of the generating model, to obtain a vector of simulated residuals *e*. In the case of the trend stationary models the residual variances were computed from the residuals from the original fitted model reported in chapter 3. The residual variance was obtained as usual under the formula:

[4.1.1] $$\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} \varepsilon_t^2}{DF}$$

where $\hat{\sigma}^2$ is the estimated residual variance, ε_t the residual value corresponding to observation "t", n the sample size and DF are the degrees of freedom in the original fitted model. The standard error of the residuals can then be obtained by simply computing:

[4.1.2] $$\hat{\sigma} = \sqrt{\hat{\sigma}^2}$$

In the case of a generating model in first differences the original model identified under the SBC was re-estimated without the drift term[3] and the standard error of the residual was thus obtained, again using [4.1.1] and [4.1.2]. Data series from a

[2] More precisely, the random numbers were generated using the rndn command in GAUSS.
[3] Obviously, where the model without drift would have been a pure random walk re-estimation without a drift term is not possible. On the other hand there are then also no coefficients which could be affected in the pure random walk series.

difference stationary data generating process were then constructed under the null hypothesis, *i.e.* without a drift term.

The simulated data series were generated using the same model specification from which the standard error of the residual had been obtained. The ARIMA (p,d,q) process was modelled by using the vector of simulated residuals *e* and the estimates for the $\phi_1 \ldots \phi_P$ and $\theta_1 \ldots \theta_q$ coefficients obtained by fitting the ARIMA (p,d,q) model, selected by SBC, to the original data series (details of the estimates are reported in appendices III.ii to III.iii). Where the selected model specified a residual process containing an autoregressive or moving average component, the residual vector was set to 200x1 with the first p values for u_{t-p} set equal to zero. Since, by the high number of rows in the vector, a hundred random values are placed before the simulated data series itself, it can then be expected that the simulated series would follow a genuine ARMA pattern by the time the 100th value has been reached. Thus, the simulated series for the trend stationary models are based on:

[4.1.3] $$p_t = a + \beta t + u_t,$$

with

[4.1.4] $$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \ldots + \phi_p u_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where ε_1 =e$[1, 1+b]$, **e** is the vector of simulated residuals, b=0 if there were no autoregressive or moving average components in the residual process (*i.e.* p=q=0) and b=100 if autoregressive or moving average components were present. An analogous procedure was employed where models were given for series in first differences, compiling the simulated series as:

[4.1.5] $p_t = p_{t-1} + v_t$

where the ε_t is specified as above, v_t follows an ARMA(p,q) process analogous to [4.1.4] above and p_1 is the first value in the original data series, so that:

[4.1.6] $p_2 = p_1 + v_2$

for the simulation of an I(1) series without drift.

4.1.2. Fitting ARIMA Models to the Generated Data Series.

Having constructed data series according to equations [4.1.3] and [4.1.5], both trend stationary and difference stationary models, were fitted to either series. In any of these cases the construction of simulated data series and subsequent model estimations were repeated 10,000 times. For those commodities where the inferred significance of the trend or drift coefficient estimate depends on stationarity assumptions as identified in chapter 3, models in levels and first differences were then fitted to the trend stationary data series generated by [4.1.3]. Table 4.1.1 below reports rejection probabilities for the null hypothesis that the trend or drift coefficient β is equal to zero as measured by the relative frequency of the rejections obtained among the 10,000 replications. In testing for the significance of $\hat{\beta}$ the asymptotic 5% critical value for a two tailed test takes a value of ±1.96. Columns two and four give the rejection rates for the relevant commodity and model, while columns three and five specify the model for which the rejection rate has been given in the preceding column.

Table 4.1.1: Probability of rejecting the null hypothesis of H_0:β=0 for simulated data series, generated from a trend stationary model.

Commodity	Proportion of Rejections	ARIMA (p,0,q)*	Proportion of Rejections	ARIMA (p,1,q)*
Beef	0.790	1,0,0	0.000	0,1,0
Cotton	0.938	1,0,0	0.346	2,1,2
Lamb	0.998	5,0,0	0.000	0,1,0
Lead	0.668	1,0,0	0.000	0,1,0
Maize	0.987	1,0,0	0.116	0,1,2
Palm Oil	0.990	1,0,1	0.000	2,1,0
Rubber	1.000	1,0,0	0.000	0,1,0
Sugar	0.983	1,0,1	0.512	0,1,2
Timber	1.000	1,0,0	0.000	0,1,0
Wheat	1.000	0,0,3	0.724	0,1,2
Wool	0.996	1,0,0	0.061	0,1,2

p:Number of autoregressive lags, q: number of moving average terms, β: trend coefficient.
* Fitted model

It should not come as a surprise that there is a high proportion of rejections of H_0 for the fitted trend stationary models since the data series are trend stationary and β is non zero by construction. What is surprising, however, are the results obtained under difference stationary model specifications. Only Sugar and Wheat have high rejection rates for the null hypothesis while Cotton, Maize and Wool have rejection rates notably in excess of zero but well below the high rejection rates obtained under the trend stationary model specification. For six of the 11 commodities, however, the proportion of rejections takes a value of 0.000. Generalising somewhat, this phenomenon seems to be particularly frequent when the trend stationary model is ARIMA(1,0,0) and the difference stationary model ARIMA(0,1,0). In those cases, in the present sample, where the percentage of rejections in difference stationary models is noticeably in excess of zero, by contrast, it appears that the error process contains a second order moving average component.

The relationship between the precise model parameterisation and the rejection rates obtained has, of course, not been investigated systematically and comprehensively here. Any generalisation should therefore be treated with caution. The particular issue of ARIMA (1,0,0) vs. ARIMA (0,1,0) model specifications is, however, treated in more detail below.

Table 4.1.2 reports the corresponding results for the models fitted to the data series generated from the difference stationary model as in [4.1.5]. Again the proportion of cases where H_0:β=0 has been rejected is reported in columns two and four, while columns three and five contain the ARIMA(p,d,q) model specifications. The last row of table 4.1.2 reports average rejection rates for those four commodities (Beef, Lead, Rubber and Timber) for which the fitted trend stationary models took an ARIMA (1,0,0) specification, while the fitted difference stationary models were ARIMA (0,1,0). Not surprisingly, the rejection rates for the ARIMA(0,1,0) models for the four commodity price series in question are rather similar and this is also the case for the rejection rates corresponding to the ARIMA(1,0,0) models fitted to data generated by a pure random walk.

Table 4.1.2: Probability of rejecting the null hypothesis of H_0: $\beta=0$ for simulated data series, generated from a difference stationary model without drift.

Commodity	Proportion of Rejections	ARIMA (p,d,q)*	Proportion of Rejections	ARIMA (p,d,q)*
Cotton	0.146	2,1,2	0.637	1,0,0
Lamb	0.051	0,1,0	0.552	5,0,0
Maize	0.178	0,1,2	0.634	1,0,0
Palm Oil	0.036	2,1,0	0.633	1,0,1
Sugar	0.156	0,1,2	0.659	1,0,1
Wheat	0.136	0,1,2	0.769	0,0,3
Wool	0.145	0,1,2	0.620	1,0,0
Beef/Lead/ Rubber/Timber	0.055	0,1,0	0.531	1,0,0

p:Number of autoregressive lags, q: number of moving average terms, β: trend coefficient.
* Fitted model.

For random walk plus drift models, the rejection rates are close to 0.05, or 5%. The rejection probability for the ARIMA(2,1,0) model in the case of Palm Oil is even lower, taking a value of 0.036. For ARMA(0,1,2) models as well as in the case of the ARIMA(2,1,2) model for Cotton, rejection probabilities are noticeably higher with values of between 0.14 and 0.18. The rejection probabilities observed for ARIMA(0,1,0) models are what one should expect for a data series where β=0 by construction, so long as the residuals are normally distributed. The rejection rates in the case of those models containing an MA(2) component are above what should be expected given the nominal 5% probability implied by the critical value. One should also note the fact that the null hypothesis is rejected in a markedly larger proportion of cases if a trend stationary model is fitted to the data series. Where the data generating process (DGP) is a pure random walk, the fitted trend stationary models reject H_0:β=0 for about 50% of the simulated series, in spite of the fact that the generating model is known not to contain a drift term. Rejection probabilities appear to be higher where the DGP has higher parameterisations even if an ARIMA(1,0,0) model is fitted in levels.

Concerning the results in table 4.1.2., it is well known that spurious rejections of the null hypothesis occur if a trend stationary model is fitted to a data series that is integrated of order one (*cf.* Newbold and Granger (1974)). More generally, the case where the generating process is a pure random walk, *i.e.*

[4.1.7] $$\Delta p_t = \varepsilon_t$$
$$p_t = p_{t-1} + \varepsilon_t$$

can be considered a special case of the form:

[4.1.8] $p_t = \eta p_{t-1} + \varepsilon_t,$

where η takes a value of one. To supplement the well documented case of spurious rejections when η=1, further simulations were conducted here, to establish whether similar results are obtained if the true value of η lies within the range of $0.7 \leq \eta \leq 0.99$. Simulated data series were obtained by generating a series of 100 normally distributed random numbers with unit variance, according to equation [4.1.8] and allowing for coefficient values of η=0.7, 0.75, 0.8, 0.85, 0.9, 0.95 and 0.99. For each of the generated series, the first value was defined as:

[4.1.9] $p_1 = \dfrac{\varepsilon_1}{\sqrt{1-\eta^2}}$

The trend stationary model

[4.1.10] $$p_t = a + \beta t + u_t, \text{ with}$$
$$u_t - \phi u_{t-1} = \varepsilon_t$$

was then fitted subsequently to the simulated series. After 10,000 replications, the percentage of times that $H_0:\beta=0$ was rejected at the nominal 5% level was obtained for each of the coefficient values used for η in the generating equation. The corresponding empirical 5% critical values for the t-ratio of the trend coefficient were obtained by ranking the observed t-ratios by numerical value, and then retaining the t-ratios corresponding to the top and bottom 2.5% of the vector. (The critical values are represented by the lowest and highest number in these 2.5% sequences if the t-ratios are sorted separately in ascending order.) The simulation results are summarised below in table 4.1.3 with rejection rates in the first row and critical values for the t-ratios in the second row. (Given the fact that the residuals used in the simulation are symmetrically distributed by construction, the absolute values of the upper and lower critical value were averaged to obtain the values reported in table 4.1.3.)

Table 4.1.3: Rejection probabilities for the null hypothesis of $H_0:\beta=0$ in an AR(1) model fitted to simulated data series generated from $p_t=\eta p_{t-1}+\varepsilon_t$

η	0.7	0.75	0.8	0.85	0.9	0.95	0.99
Rejection Rate	0.103	0.114	0.127	0.148	0.189	0.280	0.457
Critical Value	2.46	2.50	2.65	2.87	3.29	4.29	6.89

η : Autoregressive coefficient, p_t, p_{t-1} dependent and lagged dependent variable, ε_t: error term, β: trend coefficient. Values reported are absolute values of critical values.

It can be seen that the proportion of rejections of the null hypothesis increases from above 0.1 from a coefficient value of about $\eta=0.7$, and comes closer to the value of around 0.5 obtained for data generated from a pure random walk as the value of η approaches one. The absolute values of the implied critical values remain around

2.5 for autoregressive coefficients below 0.8 but then quickly approach higher values until they take values of around 6.89 as the autoregressive coefficient approaches unity.

Thus, it would appear that, even where the generating process is not strictly difference stationary, a conventional t-test for the significance of the trend coefficient can reject the null hypothesis $H_0{:}\beta=0$ too often in moderately sized samples with autoregressive processes where the value of a positive autoregressive coefficient is large.

The situation is somewhat different where the true generating model is stationary in levels around a linear trend, as was the case with the process underlying table 4.1.1, where the rejection rates for the drift term in random walk models are found to be extremely low. Some approximate calculations can be used as a basis for exploring the underlying causes of this phenomenon. One may consider a first order autoregressive stationary process with a trend coefficient β such that:

[4.1.11] $p_t = a + \beta t + u_t,\; u_t - \phi u_{t-1} = \varepsilon_t$, and $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$

Since by [4.1.11] we can write $u_t = p_t - a - \beta t$ equation [4.1.11] can alternatively be written as:

[4.1.12] $(1-\phi L)(p_t - a - \beta t) = \varepsilon_t$

First differencing [4.1.11] and defining

[4.1.13] $Z_t = p_t - p_{t-1},\, t = 2, 3, \ldots, T$

yields the model:

[4.1.14] $Z_t = \beta + v_t,\; (1-\phi L)v_t = (1-L)\varepsilon_t$

Keeping in mind that, from [4.1.14] it is true that $v_t = Z_t - \beta$ this can then be written as:

[4.1.15] $$(1-\phi L)(Z_t - \beta) = (1-L)\varepsilon_t$$

Testing H_0:β=0 at the nominal 5% level would then involve evaluating the test statistic

[4.1.16] $$\sqrt{T-1}\left|\frac{\bar{Z}}{S_Z}\right| > |1.96|,$$

where S_Z is the estimated standard error of Z, assuming wrongly that $v_t = \varepsilon_t$, so that the estimated standard error would be calculated as:

[4.1.17] $$S_Z = \sqrt{\frac{\sum_{t=2}^{T}(Z_t - \bar{Z})^2}{T-2}}\ .$$

It is known that in fact, accounting for the true structure of the residual process on the basis of [4.1.12] and [4.1.14], v_t can be written as:

[4.1.18] $$v_t = \frac{(1-L)}{(1-\phi L)}\varepsilon_t$$

It should be worthwhile then to look into the implications of the mistaken assumptions about the residual process for the estimated variance of Z_t and the more general consequences of the assumptions concerning v_t for the evaluation of the statistical significance of the estimated trend coefficient $\hat{\beta}$.

Given the sample size in the present case, it can be assumed that the asymptotic properties of the variance estimator approximately hold, so that

[4.1.19] $$S_Z^2 \rightarrow \sigma_Z^2 = \sigma_v^2$$

It is also known that the random variable $\bar{Z}$ has mean β so that, assuming normality, we have:

[4.1.20] $\bar{Z} \sim N(\beta, \sigma_{\bar{z}}^2)$,

where $\sigma_{\bar{z}}^2$ is the variance of $\bar{Z}$. Since in the present case the error terms are serially correlated, the variance of $\bar{Z}$ has to be calculated taking covariances over time into account. The variance of $\bar{Z}$ is given by:

[4.1.21]
$$\sigma_{\bar{Z}}^2 = Var\left[\frac{\sum_{t=2}^{T} Z_t}{(T-1)}\right]$$
$$= \frac{\sum_{t=2}^{T} Var(Z_t) + 2\sum_{t=2}^{T-1} Cov(Z_t Z_{t+1}) + 2\sum_{t=2}^{T-2} Cov(Z_t Z_{t+2}) + \cdots + 2Cov(Z_2 Z_T)}{(T-1)^2}$$

Since, in a stationary series, the values of covariances depend on the length of the time lag and not on the point in the time series, these summations can be rewritten as follows:

[4.1.22]
$$\sigma_{\bar{Z}}^2 = \frac{(T-1)\gamma_0}{(T-1)^2} + \frac{2(T-2)\gamma_1 + 2(T-3)\gamma_2 + \cdots + 2\gamma_{T-1}}{(T-1)^2}$$
$$= \frac{\gamma_0}{(T-1)} + \frac{2(T-2)\gamma_1 + 2(T-3)\gamma_2 + \cdots + 2\gamma_{T-1}}{(T-1)^2}$$

Following Box and Jenkins (1976) the variance for a first order autoregressive - first order moving average process is given in general by:

[4.1.23]
$$\gamma_0 = \frac{1 + \theta^2 - 2\phi\theta}{1 - \phi^2}\sigma_\varepsilon^2$$

Since in the present case θ=1, this reduces to:

[4.1.24]
$$\gamma_0 = \frac{2(1-\phi)\sigma_\varepsilon^2}{1-\phi^2}$$
$$= \frac{2(1-\phi)\sigma_\varepsilon^2}{(1-\phi)(1+\phi)}$$
$$= \frac{2\sigma_\varepsilon^2}{1+\phi}$$

It is further shown in Box and Jenkins (*op. cit.*) that for an ARMA(1,1) model in general

[4.1.25] $$\gamma_1 = \frac{(1-\theta\phi)(\phi-\theta)\sigma_\varepsilon^2}{1-\phi^2}$$

which in the present case, where θ=1 reduces to:

[4.1.26] $$\begin{aligned}\gamma_1 &= \frac{(1-\phi)(\phi-1)\sigma_\varepsilon^2}{(1-\phi)(1+\phi)} \\ &= \frac{(\phi-1)\sigma_\varepsilon^2}{1+\phi}\end{aligned}$$

and that for all γ_i with $i \geq 1$:

[4.1.27] $$\gamma_i = \phi^{i-1}\gamma_1,\ i = 1,2,\ldots,T-1$$

Substituting back into [4.1.22] then yields:

[4.1.28] $$\begin{aligned}\sigma_{\bar{Z}}^2 &= \frac{\gamma_0}{T-1} + \frac{2[(T-2)\gamma_1 + (T-3)\phi\gamma_1 + \cdots + \phi^{T-3}\gamma_1]}{(T-1)^2} \\ &= \frac{\gamma_0}{T-1} + \frac{2\gamma_1[(T-2)+(T-3)\phi + \cdots + \phi^{T-3}]}{(T-1)^2} \\ &= \frac{\gamma_0}{T-1} + \frac{2\gamma_1}{(T-1)^2}\sum_{i=0}^{T-3}(T-2-i)\phi^i\end{aligned}$$

To evaluate the probability of [4.1.16] being true one then has to evaluate:

[4.1.29] $$P\left[\frac{\sqrt{(T-1)}\,\bar{Z}}{S_Z} > 1.96\right] + P\left[\frac{\sqrt{(T-1)}\,\bar{Z}}{S_Z} < -1.96\right]$$

Focusing initially on the first term of [4.1.29] and bearing in mind that $S_{\bar{Z}}^2 \to \sigma_{\bar{Z}}^2$ and that $\sigma_{\bar{Z}}^2 = \sigma_v^2$ one can write the corresponding probability as:

[4.1.30] $$\begin{aligned}P\left[\frac{\sqrt{T-1}\,\bar{Z}}{S_Z} > 1.96\right] &\simeq P\left[\sqrt{T-1}\,\bar{Z} > 1.96\sigma_v\right] \\ &= P\left[\frac{\sqrt{T-1}\,\bar{Z} - \sqrt{T-1}\,\beta}{\sqrt{T-1}\,\sigma_{\bar{Z}}} > \frac{1.96\sigma_v - \sqrt{T-1}\,\beta}{\sqrt{T-1}\,\sigma_{\bar{Z}}}\right] \\ &= P\left[\frac{\bar{Z}-\beta}{\sigma_{\bar{Z}}} > \frac{1.96\sigma_v - \sqrt{T-1}\,\beta}{\sqrt{T-1}\,\sigma_{\bar{Z}}}\right]\end{aligned}$$

Given that $E(\bar{Z}) = \beta$ and defining:

[4.1.31] $$e = \frac{\bar{Z} - \beta}{\sigma_{\bar{Z}}}$$

it should be obvious, that with $\bar{Z}$ distributed as $\bar{Z} \sim N\left(\beta, \sigma^2_{\bar{Z}}\right)$, e is distributed as

[4.1.32] $e \sim N(0,1)$, and that

[4.1.33] $$P\left[\frac{\sqrt{T-1}\,\bar{Z}}{S_{\bar{Z}}} > 1.96\right] \simeq P\left[e > \frac{1.96\sigma_v - \sqrt{T-1}\,\beta}{\sqrt{T-1}\,\sigma_{\bar{Z}}}\right]$$

Arguing along similar lines, it can be shown furthermore, that

[4.1.34] $$P\left[\frac{\sqrt{T-1}\,\bar{Z}}{S_{\bar{Z}}} < -1.96\right] \simeq P\left[e < \frac{-1.96\sigma_v - \sqrt{T-1}\,\beta}{\sqrt{T-1}\,\sigma_{\bar{Z}}}\right],$$

where e is defined as above.

The total probability of rejecting the null hypothesis when wrongly imposing an ARIMA(0,1,0) model on a trend stationary AR(1) process is then obtained by summing the probabilities given by [4.1.33] and [4.1.34]. Table 4.1.4 below summarises the rejection probabilities obtained for different combinations of AR(1) coefficients, ϕ and drift coefficient values.

Table 4.1.4. Rejection probabilities for the t-test on the drift coefficient in differenced first order autoregressive time series.

Phi	*Beta*	P	Phi	*Beta*	P
1	*0.00*	0.050	0.8	*0.00*	0.000
	0.05	0.079		*0.05*	0.000
	0.10	0.168		*0.10*	0.000
	0.20	0.508		*0.20*	0.359
0.95	*0.00*	0.000	0.7	*0.00*	0.000
	0.05	0.001		*0.05*	0.000
	0.10	0.015		*0.10*	0.000
	0.20	0.496		*0.20*	0.233
0.9	*0.00*	0.000	0.6	*0.00*	0.000
	0.05	0.000		*0.05*	0.000
	0.10	0.001		*0.10*	0.000
	0.20	0.462		*0.20*	0.118

Calculations for rejection probabilities were made for a series with 99 observations and an error variance of one. Rejection probabilities are reported for t-tests at the nominal 5% level.

The results in Table 4.1.4 show lower rejection probabilities in the differenced model at lower autoregressive coefficient values in the generating model. As the autoregressive coefficient in the generating model falls by a comparatively small amount (from 1 to 0.95 and then to 0.9) the calculated rejection probabilities decrease dramatically. For $\phi = 0.95$ the rejection probability for a zero trend coefficient in the generating process fall from the nominal 5% level to 0.000, for a small trend coefficient value of 0.05 the rejection probability falls to 0.001 from an initial probability of 0.079. As the autoregressive coefficient falls to 0.9, the rejection rate falls to 0.000 even for a trend coefficient of $\beta = 0.05$ and takes a value of 0.001 for a trend coefficient of $\beta = 0.10$. These large falls in the rejection probabilities obtained may seem surprising. One should note then that, in spite of the rather crude nature of the above approximations, these results are well borne out by simulation evidence for comparable generating and fitted models presented below in Table 4.1.6.

This result suggests that significance tests on the drift coefficient lack power, when serial correlation in overdifferenced series is not accounted for. This is also reflected in the rejection probabilities calculated for Beef, Lead, Rubber and Timber, the four commodities in table 4.1.1. where the alternative trend and difference stationary representations are ARIMA (1,0,0) and ARIMA (0,1,0) respectively. In all four cases, the rejection probabilities calculated on the basis of the estimated model parameters and using [4.1.33] and [4.1.34], take a value of zero at the four digit level. The trend coefficients for these commodity price series can be normalised by dividing them by the standard error of the residual. A direct comparison with the theoretical results in table 4.1.4 (and the simulation results presented later on in table 4.1.6.) is then possible by making reference to both the normalised trend coefficient and the estimated autoregressive coefficient.

Table 4.1.5. estimated AR(1) and normalised trend coefficients for selected commodities.

Commodity	AR(1)	Trend Coefficient
Beef	0.910	0.067
Lead	0.800	-0.033
Rubber	0.800	-0.104
Timber	0.680	0.076

The low rejection probabilities in table 4.1.6. below correspond to what one should expect from the predictions in table 4.1.4., and the rejection probabilities predicted from tables 4.1.4 and 4.1.5 are consistent with the results in table 4.1.1.

Estimation and test results may improve if serial correlation in the form of autoregressive as well as moving average processes is allowed for in difference stationary models, through a more elaborate parameterisation. This would also make the result in table 4.1.1. appear plausible, where higher rejection rates were associated with ARIMA models which contained moving average terms. Table 4.1.6. gives the results obtained when first order autoregressive trend stationary series are generated by Monte Carlo Simulation and difference stationary models of type ARIMA(0,1,0) and ARIMA(1,1,1) fitted to the generated series subsequently. The trend stationary model simulated was of the form:

[4.1.35] $$y_t = a + \beta t + u_t,\ u_t - \phi u_{t-1} = \varepsilon_t$$

with the trend coefficient taking values between 0 and 0.2 while the values for the AR(1) coefficient were confined to the range 0.6-1. For each simulated model, 5,000 replications were used to obtain the rejection rates for $H_0 : \beta = 0$ evaluated at the asymptotic critical value of +/-1.96. The results are given in Table 4.1.6. below.

Table 4.1.6. Empirical rejection rates for trend and drift coefficients evaluated at the nominal 5% critical level.

ϕ	β	ARIMA(1,0,0)	ARIMA(0,1,0)	ARIMA(1,1,1)
1	**0.00**	0.528	0.049	0.107
	0.05	0.580	0.083	0.142
	0.10	0.673	0.185	0.248
	0.20	0.895	0.511	0.543
0.95	***0.00***	*0.265*	*0.000*	*0.028*
	0.05	*0.484*	*0.001*	*0.069*
	0.10	*0.834*	*0.017*	*0.172*
	0.20	*0.997*	*0.496*	*0.612*
0.90	**0.00**	0.186	0.000	0.025
	0.05	0.643	0.000	0.118
	0.10	0.972	0.002	0.301
	0.20	1.000	0.465	0.711
0.80	***0.00***	*0.120*	*0.000*	*0.022*
	0.05	*0.925*	*0.000*	*0.352*
	0.10	*1.000*	*0.000*	*0.648*
	0.20	*1.000*	*0.398*	*0.893*
0.70	**0.00**	0.098	0.000	0.022
	0.05	0.997	0.000	0.645
	0.10	1.000	0.000	0.871
	0.20	1.000	0.294	0.972
0.60	***0.00***	*0.087*	*0.000*	*0.000*
	0.05	*1.000*	*0.000*	*0.853*
	0.10	*1.000*	*0.000*	*0.964*
	0.20	*1.000*	*0.216*	*0.998*

ϕ: value of the autoregressive coefficient, β: value of the trend coefficient in the data generating process. The null hypothesis tested is $H_0 : \beta = 0$.

It is apparent from table 4.1.6. that, at least for low values of the autoregressive coefficient, the rejection rates for the null hypothesis of a zero trend coefficient are high, often close to one for both ARIMA (1,0,0) and ARIMA (1,1,1) models. These rejection rates fall as the value of the autoregressive coefficient approaches one and also for lower values of the trend coefficient in the data generating process. For those data series which were generated under the null hypothesis, *i.e.* with β=0, it is also apparent that the observed rejection rates increase as the value of the autoregressive coefficient approaches 1. This again confirms the results presented in table 4.1.3. It can be seen throughout that the rejection rates for the ARIMA(1,1,1) model are consistently above those for the simple random walk with drift specification. There is, however, some danger of spurious rejection of the null hypothesis when the ARIMA(1,1,1) specification is used since the rejection rates obtained for I(1) series under the null still amount to 10.7% when testing at the nominal 5% level. On the other hand, rejection rates for ARIMA (1,1,1) models are still noticeably lower than for the corresponding trend stationary model for large values of the autoregressive coefficient. T-tests on the drift coefficient estimate in ARIMA (1,1,1) models can thus be seen to be more powerful than in the simple random walk plus drift specification. Although the probability of wrongly rejecting the null hypothesis increases somewhat, the ARIMA(1,1,1) model specification seems to offer a better balance between the risk of incurring Type I or Type II errors[4] compared to the simple random walk plus drift model as an alternative to the trend stationary model specification. Type II errors remain a problem at low drift coefficient values and for highly persistent serial correlation. Even for near integrated series though, there are substantial improvements over the random walk plus drift alternative.

It thus seems desirable to obtain a higher parameterisation for the difference stationary models fitted to the commodity price series in the sample. The above scenario of ARIMA(1,0,0) models being selected by SBC for the model in levels while the difference stationary minimum SBC model is a simple random walk with

[4] A rejection of the null hypothesis when it is in fact true is referred to as Type I error, whereas the term Type II error refers to the possibility of accepting the null hypothesis when it is false.

drift is given in ten cases. Table 4.1.7. shows the results of an attempt to correct selection by SBC for underparameterisation. Columns two and three give the drift coefficient estimates and t-ratios for ARIMA(p,1,q) models which have been selected by minimum SBC subject to the constraints that $p+q \leq 5, p \geq 1$ and $q \geq 1$. The values in column four (labelled t-ratios (SBC)) are the t-ratios obtained for the original model selected by SBC without constraining p and q as in the present case. Column five gives the new minimum SBC model for the difference stationary case. The additional constraints limit the most parsimonious model selection possible to ARIMA(1,1,1) in order to avoid the low power observed in the case of simple random walk plus drift models[5].

Table 4.1.7. Drift Coefficients and t-ratios for higher parametarisations of the Difference Stationary Model Selected by SBC.

Commodity	Drift coefficient	t-ratio	t-ratio (SBC)	ARIMA (p,1,q)
Coffee	0.005	0.380	0.093	1,1,1
Tea†	-0.011	-1.113	-0.573	1,1,2
Beef	0.008	0.365	0.387	1,1,1
Lamb	0.016	0.722	0.715	1,1,1
Bananas	0.000	0.044	0.043	1,1,1
Tobacco†	0.003	0.268	0.218	3,1,2
Rubber	-0.029	-6.540	-1.041	1,1,2
Timber	0.011	1.942	0.493	1,1,1
Copper	-0.008	-0.702	-0.490	1,1,2
Tin	-0.004	-0.178	-0.172	1,1,1
Lead	-0.005	-0.917	-0.414	1,1,1

† Data series from 1900-1997 only. Minimum SBC models have been constrained to p+q<=5, p,q >=1.

[5] It will be observed that the models selected for Tea, Tobacco and Copper by SBC using the additional constraints are not ARIMA(1,1,1) models. However when ARIMA(1,1,1) models are fitted, they do not indicate the presence of significant drift terms either. The t-ratios on the drift coefficient estimate are -0.111, 0.180 and -0.457 for Tea, Tobacco and Copper respectively *Cf.* Appendix IV.i).

Imposing the above constraint on the minimum SBC selection, it appears that Rubber now has a statistically significant trend coefficient, while this was not so under the pure random walk plus drift specification. The drift coefficient of another commodity, Timber, now appears significant at the 10% level, and the t-ratios for the coefficients of other price series (Coffee, Tea, Copper and Lead) have increased somewhat, but not sufficiently for the coefficients to be significant at the 5% or 10% levels. The t-ratios on the remaining drift coefficients in table 4.1.7. have changed little for the higher SBC specification.

Strictly speaking, the above simulation results only demonstrate the empirical superiority of ARIMA(1,1,1) models over the alternative of ARIMA(0,1,0) models when the data generating process is ARIMA(1,0,0). It remains to be seen how more elaborate models perform in general. To allow for less parsimonious model specifications for a number of data generating processes, the Akaike Information Criterion can be used, instead of the SBC employed so far, to allow more generally for the possibility of underparameterised difference stationary models. Since the SBC is known to select more parsimonious model specifications than the AIC, model selection by AIC may in the present case serve to safeguard against selecting underparameterised difference stationary models and thus avoid the associated problems of low powered hypothesis tests. The coefficient estimates and t-ratios from models selected by AIC are listed in columns two and three of table 4.1.8. The t-ratios corresponding to the minimum SBC models reported in chapter 3 are reported in column four, while the model specification for selection by minimum Akaike Information Criterion is reported in column five. (Details for the estimation results underlying Tables 4.1.7. and 4.1.8. are given in appendix IV.i.).

Table 4.1.8: Drift coefficients for ARIMA(p,1,q) models selected by AIC.

Commodity	Drift / minimum AIC	t-ratio	t-ratio (SBC)	ARIMA (p,1,q)
Coffee	0.003	0.154	0.093	0,1,2
Cocoa	-0.009	-0.457	-0.457	2,1,0#
Tea†	-0.011	-1.017	-0.573	0,1,2
Rice	-0.012	-2.423*	-2.423*	1,1,2#
Wheat	-0.011	-1.923	-1.154	0,1,4
Maize	-0.010	-1.341	-1.341	0,1,2#
Sugar	-0.011	-3.322*	-0.844	0,1,5
Beef	0.008	0.387	0.387	0,1,0#
Lamb	0.015	0.585	0.715	4,1,1
Bananas	0.000	0.043	0.043	0,1,0#
Palm Oil	-0.010	-1.119	-0.429	0,1,3
Cotton	-0.007	-0.641	-0.765	2,1,3
Jute	-0.008	-0.665	-0.665	0,1,2#
Wool	-0.014	-1.850	-1.850	0,1,2#
Tobacco†	0.003	0.268	0.189	3,1,2
Rubber	-0.029	-6.540*	-1.041	1,1,2
Timber	0.012	8.297*	0.493	0,1,5
Copper	-0.008	-0.702	-0.490	1,1,2
Aluminium	-0.019	-7.931*	-7.931*	1,1,2#
Tin	-0.003	-0.172	-0.172	0,1,0#
Silver	-0.003	-0.229	-0.229	2,1,0#
Lead	-0.008	-1.158	-0.414	0,1,4
Zinc	0.000	0.184	0.184	1,1,2#

† Data Series from 1900-1997 only. * Significant coefficient at the 5% level. # Model selected does not change when using AIC instead of SBC.

For a number of commodities (Cocoa, Rice, Maize, Beef, Bananas, Jute, Wool, Aluminium, Tin, Silver and Zinc), the selected model does not change when the AIC rather than the SBC is used for model selection. Of the drift coefficient estimates shown, five: those for Rice, Sugar, Rubber, Timber and Aluminium, now appear to be statistically different from zero. Of these, the coefficients for

Aluminium and Rice appeared to be significant previously, when models where selected by SBC (and the same models were selected). In the cases of Rubber, Timber and Sugar, selecting a more elaborately parameterised model does indeed change the t-statistic on the drift coefficient. For a number of other commodities (Coffee, Tea, Wheat, Palm Oil, Tobacco, Copper and Lead), the t-ratio on the coefficient estimate obtained increases in absolute value but this change is not sufficient to affect the results of the hypothesis test on the drift coefficient at the 5% level. In the case of Wheat, the t-ratio on the drift coefficient estimate falls from -1.154 to -1.923 which is just above the asymptotic 5% critical value of -1.96. Given that the fall in the t-ratio is observed after introducing a more elaborate model parameterisation and given that it is very close to the required critical value (as well as being clearly below the 10% critical value of -1.65) the drift coefficient obtained when selecting the model for Wheat by AIC[6] is considered significant. In the cases of Lamb and Cotton the absolute value of the t-ratio actually falls when the model in first differences is re-selected using the AIC. The assumption that evidence for a trend has been obscured because the absolute values of the t-statistics obtained have been understated by excessively parsimonious model selection therefore receives empirical support only for the price series of Sugar, Rubber, Timber and Wheat.

---

[6] In view of the relatively high rejection rates obtained for Wheat in Table 4.1.1., one may ask, how much importance should be attached to the higher t-ratio recorded when re-selecting by AIC. Additional simulations -following the above methodology- show that the rejection rate recorded for the model selected by AIC is noticeably higher at 0.980. The probability of spurious rejections is also somewhat higher at 0.184.

4.2 A Priori Inference on Unit Roots on the Basis of a Stationarity Test.

The discrepancies in the inference about the presence of a significant trend term depending on the model fitted underline the importance of distinguishing trend and difference stationary processes. It has been shown above how the reliability of the hypothesis test can be improved if the difference stationary model is more adequately specified. In the case of Lead, the ARIMA(1,1,1) model estimated as an alternative to the random walk model selected by SBC shows signs of overdifferencing, since the estimate of the moving average parameter is on the invertibility boundary. Against this background, a comparison of the unit root test results with the results from a stationarity test appears to be of interest. While a fully satisfactory solution to the problem of pre-testing is not available for the present case, there are some problems which can be accounted for if a different testing procedure is used.

It is known that Dickey-Fuller type tests tend to reject the null hypothesis of a unit root too frequently in the presence of a moving average component where the absolute value of the moving average coefficient is at or near to the invertibility boundary (*cf.* Agiakloglou and Newbold (1992)). The test employed to allow for this possibility in the present case is the Leybourne McCabe test, a stationarity test which is reported to be robust to the presence of a large moving average coefficient as well as to overfitting of autoregressive lags (Leybourne McCabe (1994)).

The Leybourne McCabe test is a stationarity test which, like the test developed by Kwiatkowski *et.al.* (1992), tests the null hypothesis of stationarity against the alternative hypothesis of a unit root, by modelling a time series as the sum of a deterministic trend, a random walk and a white noise residual term (Kwiatkowski *et. al.* (*op. cit.*)).

The testing procedure used here follows Leybourne and McCabe (1994) and uses the modified variance estimator and implementation procedure outlined by Leybourne and McCabe (1999). The time series in levels is modelled as:

[4.2.1.] $\Phi(L)p_t = a + \beta t + a_t + \varepsilon_t$

with, $a_t = a_{t-1} + \eta_t$

the random walk component, βt the deterministic trend and coefficient as before, α a constant and ε_t a white noise error term. The dependent variable is p_t, the relative price of the relevant primary commodity in natural logarithms. Given $a_0 = 0$ and $\eta \sim iid(0, \sigma_\eta^2)$ it follows that the null hypothesis of stationarity implies that $\sigma_\eta^2 = 0$, so that $a_t = 0$. The above time series model can be represented in first differences as:

[4.2.2.] $\Phi(L)(1-L)p_t = \beta + (1-\theta L)\zeta_t$

from where it can be seen that the null hypothesis $H_0 : \sigma_\eta^2 = 0$ implies that p_t is I(0) and $\theta = 1$, while it follows from $H_1 : \sigma_\eta^2 > 0$ that $\theta < 1$. Since for a (trend) stationary series equation [4.2.2.] would be overdifferenced, it is expected that the moving average coefficient in the estimated model would have a value close to one (see also [4.1.14 *ff.*]), and vice versa that it takes a value below unity if the data generating process is not either stationary or trend stationary in levels.

To test for the trend stationarity of an autoregressive time series, allowing for p autoregressive terms and a first order moving average term, the following equation was estimated by exact maximum likelihood estimation, using the GAUSS[7] time series package (*cf.* Leybourne and McCabe (1999) for the implementation of the testing procedure):

[4.2.3.] $$\Delta p_t = a + \phi_1 \Delta p_{t-1} + \phi_2 \Delta p_{t-2} + \ldots + \phi_p \Delta p_{t-p} + \zeta_t - \theta \zeta_{t-1}$$

from which σ_ε^2 in the levels equation can be inferred from: $\tilde{\sigma}_\varepsilon^2 = \hat{\sigma}_\zeta^2 \hat{\theta}$, *i.e.* as the product of the maximum likelihood estimate of the moving average parameter and the residual variance from the testing equation (Leybourne and McCabe (1999)). The number of autoregressive parameters was chosen through general to specific testing starting with p=5 and testing at the 10% significance level. Instead of the t-ratio on the p^{th} autoregressive coefficient[8] the Z statistic $Z = \hat{\phi}_p \hat{\theta} \sqrt{n}$, with $\hat{\phi}_p$ the autoregressive coefficient estimate on the p^{th} lag, $\hat{\theta}$ the moving average coefficient estimate and n the number of observations in the differenced series. Following Leybourne and McCabe (1999), this should identify the true autoregressive order of the generating process so long as $p \geqq p$, where p is the starting value of the number of lagged differenced terms and p the true number of autoregressive lags. After estimating equation [4.2.3], the original data series were filtered using the estimated ϕ_i coefficients as in:

[4.2.4.] $$p_t^* = p_t - \sum_{i=1}^{p} \hat{\phi}_i p_{t-i}$$

The least squares residual vector $\hat{\varepsilon}_t^* = p_t^* - \hat{a} - \hat{\beta} t$ was then obtained by first regressing p_t* on a constant and linear trend using OLS and subtracting the

---

[7] The GAUSS time series package computes the autoregressive terms for the residual rather than for lagged dependent variables. The estimates for the ARMA coefficients should be identical though.

[8] Leybourne and McCabe highlight the fact that there are potential problems in computing standard errors when the moving average coefficient is at or near the invertibility boundary.

predicted from the filtered values. In order to obtain the Leybourne McCabe test statistic, a further matrix, V, was computed such that the i,j[th] element of this matrix takes value i if i<j, value j if j<i and value i=j if i=j. Given subsequent pre- and post multiplication with the least squares residual vector ε, the matrix V needs to be symmetric, so that a TxT version of matrix V would take the form:

[4.2.5] $$V=\begin{vmatrix} 1 & 1 & (\ldots) & 1 \\ 1 & 2 & (\ldots) & 2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 2 & (\ldots) & T\text{-}1 \\ 1 & 2 & (\ldots) & T \end{vmatrix}$$

Given V, $\hat{\varepsilon}^*$, and $\tilde{\sigma}_\varepsilon^2$, where $\tilde{\sigma}_\varepsilon^2$ is computed as specified above, the Leybourne-McCabe test statistic for the test including a trend, $\tilde{s}_\beta(p)$, can then be obtained as:

[4.2.6] $$\tilde{s}_\beta(p)=\frac{\hat{\varepsilon}^{*\prime}V\hat{\varepsilon}^*}{\tilde{\sigma}_\varepsilon^2 T^2},$$

where p is the number of autoregressive lags in the testing equation. The general to specific methodology adopted to determine the number of autoregressive lags in the testing equation above asymptotically identifies the correct number of autoregressive lags only if the maximum number of lags is at least as large as the true number of autoregressive lags in the data generating process. For this reason, it appears desirable to arrive at any empirically determined number of autoregressive lags from an initial higher number of lags in the testing equation. Since in the present study, general to specific testing was undertaken from an imputed maximum number of 5 lags, tests with a higher maximum number of lags were undertaken for all those commodities, were the final number of autoregressive lags was found to be 5. Testing from a maximum number of 8 lags, this yielded different final results only in the case of Rice, where now a total number of eight AR lags appears significant[9]. (Allowing for five lags only, the moving average coefficient for Rice appeared to be on the invertibility boundary and the Leybourne-McCabe test statistic took a value of 0.0323, so that the stationarity hypothesis could not be rejected. With a higher number of autoregressive lags, the coefficient on the moving average term is no longer equal to one though it still takes a value of 0.791. The null hypothesis of trend stationarity can be rejected in this case.)

The test statistics obtained together with the values and significance levels of the moving average parameter and the number of autoregressive lags in the testing equation are reported in table 4.2.1 below. Kwiatkowski *et. al.* (1992) report an upper 5% critical value of 0.146 for this type of stationarity test when a trend term is included (A fuller listing of the critical values and the full specifications for the final testing equations are given in appendix IV.ii.)

It has been pointed out by Ahrens and Vijaya (1997) that ordinary ADF tests can perform sufficiently well if the coefficient on the moving average component is not significantly different from zero. The Leybourne McCabe test furthermore is based on the assumption that the moving average coefficient is constrained to $0 < \theta \leq 1$ (*cf.* Leybourne and McCabe (1999)) while a number of the estimated moving average coefficients reported below have negative signs. Thus the Leybourne McCabe test results are taken into consideration only for those cases where the

[9] These eight autoregressive lags were selected again when general to specific testing for the series for Rice was performed from a maximum number of 10 autoregressive lags.

estimated moving average coefficient was significantly different from zero as well as positive. (Indeed, in those cases where the moving average coefficient takes a negative value, the estimated residual variance $\tilde{\sigma}_{\varepsilon}^{2}$ would, somewhat nonsensically, take a negative value, since it is inferred on the basis of the moving average coefficient. This would, moreover result in a negative value for the overall test statistic.)

Based on the above, the null hypothesis of stationarity can be rejected in five cases: Rice, Maize, Cotton, Jute and Wool. In a further eight cases (Coffee, Wheat, Lamb, Palm Oil, Rubber, Aluminium, Lead and Zinc) the null hypothesis can not be rejected. In all these cases, the estimated moving average coefficient is on the invertibility boundary, so that no reliable estimates can be obtained for the t-ratios. (One therefore may proceed on the assumption that the coefficient on the MA statistic be statistically significant if the Leybourne McCabe test is to be applied in these cases. Since an MA coefficient close to one is expected in the case of stationary time series this assumption is maintained for estimated coefficient values on the invertibility boundary.) In all the remaining cases, the estimated moving average coefficients take values below zero and are -with the exception of the price series for silver- not significantly different from zero. Leybourne and McCabe (1999) explicitly formulate their model for the case where $0 < \theta \leq 1$ so that, for negative values of θ, the test would not be applicable in any case.

Table 4.2.1. Stationarity Test Results for the Leybourne McCabe Test

Commodity	$\tilde{S}_{\beta}$	MA	t-MA	p
Coffee	0.035†	1	(...)	1
Cocoa	-11.237	-1	(...)	1
Tea	-25.028	-1	(...)	1
Rice	0.520	0.791	4.591	8
Wheat	0.043†	1	(...)	5
Maize	0.884	0.779	6.715*	2
Sugar	-5.262	-0.592	0.318	5
Beef	-46.940	-0.055	-0..535	0
Lamb	0.030†	1	(...)	5
Bananas	-35.573	-0.538	-0.871	1
Palm Oil	0.025†	1	(...)	3
Cotton	0.610	0.783	4.428*	1
Jute	0.740	0.818	5.328*	1
Wool	1.101	0.817	6.768*	1
Tobacco	-46.253	-0.706	-1.943	5
Rubber	0.033†	1	(...)	1
Timber	-1.144	-1	(...)	3
Copper	-9.483	-0.504	-0.716	1
Aluminium	0.025†	1	(...)	3
Tin	-40.747	-0.066	-0.644	0
Silver	-27.148	-0.968	-10.666*	3
Lead	0.088†	1	(...)	1
Zinc	0.084†	1	(...)	2

$\tilde{S}_{\beta}$: Leybourne McCabe Test Statistic, MA: estimated coefficient on the moving average parameter, t-MA: the t-ratio for the estimated coefficient on the moving average parameter, p: autoregressive order underlying the Leybourne McCabe test.
† Value below the critical value of the test statistic for rejection of the null hypothesis of stationarity.
* Statistically significant coefficient for the moving average coefficient at the 5% level.
(...) t-ratio is not reported since the MA root is on the boundary.

In the cases of Coffee and Lead[10], where there was evidence of overdifferencing when fitting an ARIMA (1,1,1) model, the Leybourne McCabe test does indeed indicate stationarity where the ADF test did not. For several of the remaining price series the ADF and Leybourne McCabe test also lead to opposite conclusions.

[10]The same could be said for Rubber when considering the ARIMA(1,1,1) model estimated for the Leybourne McCabe test. In this model too, the estimate of the moving average coefficient is on the invertibility boundary.

Table 4.2.2 below summarises the different stationarity conclusions from the Augmented Dickey Fuller test and the Leybourne McCabe test.

Table 4.2.2. Conclusions on the Order of Integration Based on Unit Root and Stationarity Tests

Commodity	ADF	LMC
Coffee	**DS**	**TS**
Rice	DS	DS
Wheat	**DS**	**TS**
Maize	DS	DS
Lamb	TS	TS
Palm Oil	**DS**	**TS**
Cotton	DS	DS
Jute	DS	DS
Wool	DS	DS
Rubber	**DS**	**TS**
Aluminium	TS	TS
Lead	**DS**	**TS**
Zinc	TS	TS

ADF: Augmented Dickey Fuller Test,
LMC: Leybourne McCabe Test
DS: Difference Stationary, TS: Trend Stationary

In the cases of Coffee, Wheat, Palm Oil, Rubber and Lead, the inference of trend stationarity from the Leybourne McCabe test is in opposition to the stationarity conclusions that would be obtained from the Augmented Dickey-Fuller test with trend. The series for Lamb, Aluminium and Zinc are identified as trend stationary by both tests while either test identifies the price series for Rice, Maize, Cotton, Jute and Wool as difference stationary. Among those series identified as trend stationary by the ADF test, the ones for Sugar and Timber are the only ones to which the Leybourne McCabe test is not applicable since the estimated moving average coefficient is negative.

If one were to follow a strategy of pre-testing, the Augmented Dickey Fuller test could be applied in those cases where the estimated coefficient on the moving average term in the Leybourne McCabe testing equation is insignificant or negative while one may use the Leybourne McCabe test, where the estimated coefficient on the moving average term is either significant or on the boundary of the invertibility region. Proceeding thus, the commodity price series that would be classified as trend stationary would be Coffee, Wheat, Sugar, Lamb, Palm Oil, Rubber, Timber, Aluminium, Lead and Zinc. The remaining commodities (Cocoa, Tea, Maize, Beef, Banana, Cotton, Jute, Wool, Tobacco, Copper, Tin and Silver) would be classified as difference stationary time series.

The Leybourne McCabe test can overcome the problems which the ADF test is subject to if there are large moving average coefficients in the testing equation, so long as the number of autoregressive lags is not underspecified. Other problems associated with pre-testing such as the possible impact of structural breaks or the appropriate specification of the null hypothesis and of significance levels in testing procedures do, however, remain. A further problem suggested by the above simulation results is that simple t-tests on trend coefficient estimates can still be subject to small sample distortions from large positive autoregressive components. In the following section, the pre-test results obtained so far are combined with a number of methods which aim at accounting for the impact of first order autocorrelation in the data generating process. A modified approach to pre-testing is also used.

4.3 Modified Testing Methods for a Deterministic Trend in First Order Autoregressive Time Series

Various methods of testing for the presence of a deterministic trend in first order autoregressive time series have been used in a study by Sun and Pantula (1999). A number of these will be employed here to assess the evidence in favour of a deterministic trend in those of the present data series where an AR(1) representation was found to be adequate for the data series in levels, and where the t-test statistic identifies the trend coefficient estimate as statistically significant at asymptotic critical values. Sun and Pantula use a number of estimation methods such as OLS, Maximum Likelihood estimation and Conditional Maximum Likelihood to estimate the trend coefficient and assess the evidence in favour of its statistical significance, relying on previous inferences on the presence of unit roots. They then proceed to obtain simulation evidence on the adequate critical values after adjusting the value of the estimated autoregressive coefficient for an imputed estimation bias. The authors applied all estimation and testing methods to the original data series as well as to data series adjusted for first order autocorrelation. In the present study, some of these methods were applied to the commodity price series used, employing OLS and Maximum Likelihood estimation only. The data series used are those for which an ARIMA (1,0,0) model has been selected by SBC in chapter 3. Assumptions on stationarity are formed on the basis of the ADF and Leybourne McCabe test results reported above.

4.3.1. OLS with adjusted t-ratios

The first variety of what Sun and Pantula refer to as the pre-test method is based on calculating initially an adjusted t-statistic which is then compared with the appropriate critical values depending on whether or not the time series in question has been identified as trend stationary or difference stationary. Sun and Pantula start from a trend stationary model with first order autoregressive residuals of the form:

[4.3.1.] $$Y_t = a + \beta t + u_t, \text{ and } u_t = \phi u_{t-1} + \varepsilon_t$$

which is clearly equivalent to the AR(1) model considered here. The statistical significance of the estimated trend coefficient is then assessed on the basis of the adjusted t-statistic:

[4.3.2.] $$t^*_{OLS} = \left[\sqrt{\frac{1-\hat{\phi}_{OLS}}{1+\hat{\phi}_{OLS}}}\right] t_{OLS},$$

where $\hat{\phi}$ is the OLS estimate of the first order autoregressive coefficient and t_{OLS} is the conventional, unadjusted t-statistic (*cf.* Sun and Pantula (*op. cit.*)). For the case where ϕ=1, Sun and Pantula obtained, by simulation, a critical value of +/-7.65 for 100 observations. In the case where ϕ<1, the conventional critical value of +/-1.96 was deemed appropriate by Sun and Pantula.

In the present case, simulations were conducted to verify the critical value for the unit root case, *i.e.* data series were generated for the model in equation [4.3.1.] under the null hypothesis that ϕ=1 and $\beta = 0$. The autoregressive residual series were constructed using normally distributed random numbers for 99 observations (again following the basic methodology outlined in section 4.1.1.). The full trend stationary model [4.3.1.] was then fitted to the generated series and the t*-ratios for the estimated coefficient $\widehat{\beta}$ were retained over 10,000 replications. Averaging the top and bottom 2.5% values for the t*-ratios of five simulation runs yielded an empirical critical value of +/-7.682. This value appears to be reasonably close to the critical values of +/-7.88 for 50 observations and +/-7.65 for 100 observations obtained by Sun and Pantula (who also used 10,000 replications in this case), so that the ±7.682 critical value obtained here is taken as representative for the present case.

In the light of the above findings on the properties of the asymptotic t-statistic in moderately sized samples, the conventional 5% critical value of ±1.96 has not been considered appropriate in the present case. To evaluate the implications of large autoregressive coefficient values in terms of spurious rejections of the null hypothesis and in order to compute the implied critical values, simulations[11] equivalent to the ones presented in table 4.1.3. for Maximum Likelihood estimation were conducted estimating 4.1.8 by OLS rather than Maximum Likelihood and correcting the t-ratio for the trend coefficient estimate as in [4.3.2.] above. The results obtained for data series with 100 observations are reported in table 4.3.1. below. Critical values were obtained as before by retaining the top and bottom 250TH t-ratio from the series of 10,000 t-ratios after these had been sorted in ascending order. Since the simulated random errors are again normally distributed and therefore symmetric by construction, the upper and lower critical values have been averaged to yield the values reported in the second row of table 4.3.1.

[11] Here and in the remainder of the present chapter, the methodology adopted for simulating integrated or correlated residual series generally follows the methodology outlined in section 4.1.

Table 4.3.1: Rejection probabilities for the null hypothesis of H_0:β=0 in an OLS model with AR(1) errors fitted to simulated data series generated from $p_t=\eta p_{t-1}+\varepsilon_t$

η	0.7	0.75	0.8	0.85	0.9	0.95	0.99
Rejection Rate	0.086	0.095	0.107	0.122	0.149	0.199	0.337
Critical Value	2.31	2.39	2.50	2.67	2.93	3.45	5.42

η : Autoregressive coefficient, p_t, p_{t-1} dependent and lagged dependent variable, ε_t: error term, β: trend coefficient. The reported critical values are absolute values.

As can be seen from table 4.3.1. rejection rates and implied absolute critical values for the trend coefficient are somewhat lower in the OLS than in the Maximum Likelihood case for the given sample size.

The results obtained are close to simulation results on the basis of estimated coefficients from fitting first order autoregressive models for the three trend stationary commodity price series to the original data series by OLS. (See Table 4.3.2. below.) Here the empirical critical values for the trend stationary case were obtained by first estimating $\hat{\phi}$ in the AR(1) model fitted to the original data series: as in Sun and Pantula (1999) the dependent variable was first regressed on a trend and constant. The residuals from this regression were then regressed on their lagged values to obtain $\hat{\phi}$. It is worth bearing in mind that, as should be apparent from the results above, the inferred critical values depend on the magnitude of the autoregressive coefficient in the original data generating process, and that underestimation of the AR(1) coefficient will in turn lead to understated absolute critical values in the simulation results. Sun and Pantula (*op. cit.*) state that estimates of the autoregressive coefficient are subject to downward bias regardless of whether OLS or Maximum Likelihood estimation is used (a suggested method for bias correction is explored below).

To obtain the lower and upper critical values for the commodities in question, simulations were performed under the null hypothesis of a zero trend coefficient for 99 observations and over 10,000 replications. Simulations were run five times for each commodity and the absolute values of the 5% critical values obtained were averaged subsequently. The absolute critical values obtained for each commodity price series and the estimated autoregressive parameters used in the simulation are reported in table 4.3.2 below. (Upper and lower critical values have again been averaged. As the residuals used in the simulation are normally distributed by construction here and throughout the remainder of the section, this will also be done for later simulation results.)

Table 4.3.2. Empirical Critical Values and AR(1) Coefficients for Selected Commodity Price Series

Commodity	CV	$\hat{\phi}$
Rubber	2.546	0.803
Timber	2.328	0.685
Lead	2.544	0.799

CV: Critical Value, $\hat{\phi}$: estimated autoregressive coefficient, The reported critical values are absolute values.

The critical values in Table 4.3.2. increase consistently with the value of the autoregressive coefficient in the data generating process. This is very much in line with the above findings on empirical rejection rates for near integrated first order autoregressive processes presented in tables 4.1.3 and 4.3.1.

This complication should be borne in mind when any inference on the significance of the trend coefficient is attempted. One can, however, venture some preliminary conclusions at this stage. After pre-testing for stationarity using the ADF test and Leybourne McCabe test as described above, only two commodities identified as trend stationary (Rubber and Timber) are shown to have statistically significant trend terms. Another trend stationary commodity price series (for Lead) and all of the difference stationary commodity price series with fitted first order autoregressive models have trend terms which are statistically insignificant. The estimated trend coefficients together with their t-ratios and, in those cases where the estimated autoregressive coefficient takes a value below unity, the simulated OLS critical value are reported in table 4.3.3. below,

Table 4.3.3. Estimated Trend Coefficients, t-ratios and Empirical 5% Critical Values for Selected Primary Commodity Price Series.

Commodity	Trend ($\hat{\beta}$)	t-ratio* for $\hat{\beta}$	Critical Value
Beef	0.017	2.242	7.682
Cotton	-0.010	-3.416	7.682
Lead	-0.005	-1.792	2,544
Maize	-0.010	-4.476	7.682
Rubber	-0.027	-5.728	2,546
Timber	0.011	7.222	2,328
Wool	-0.016	-4.664	7.682

$\hat{\beta}$: Estimated Trend Coefficient, *t-ratios are adjusted for first order serial correlation. The reported critical values are absolute values.

Had inferences on trend stationarity and difference stationarity been made on the basis of ADF tests alone, Rubber would have been classified as difference stationary, and its trend coefficient estimates would have been classified as statistically insignificant at the reported critical value of +/-7.682. It is also worth noting, however that the t-ratio for Timber -which takes a value of 7.222- is close to the reported 5% critical value for the I(1) case.

4.3.2. Maximum Likelihood Estimation Results

The above estimations were repeated using Maximum Likelihood estimation instead of OLS. Sun and Pantula (*op. cit.*) re-express the equation for the AR(1) model [4.3.1.] to yield:

[4.3.3.] $$Y_t = b_0 + b_1 t + \phi Y_{t-1} + \varepsilon_t$$

with t=2,3,...T, $b_0 = a(1-\phi)+\beta\phi$ and $b_1 = (1-\phi)\beta$. ϕ is identical to the autoregressive parameter in the specification of the AR(1) residual process u_t. It is worth noting moreover, that when setting $\phi = 1$, both [4.3.1.] and [4.3.3.] will reduce to a random walk with drift.

In the present study, Maximum Likelihood estimations were again performed for the original model [4.3.1.]. Critical values for the null hypothesis of a zero trend coefficient for an I(1) price series were obtained by simulation. Again, normally distributed random numbers were used to construct random walk residual series for 99 observations and over 10,000 replications following the methodology outlined above. The average lower and upper 2.5% critical value for the t-ratios obtained by Maximum Likelihood estimation is +/-8.319, compared with a value of +/-7.55 quoted by Sun and Pantula (*op. cit.*).

The observed critical values from simulations conducted for the trend stationary price series, using the same methodology as above, are given in table 4.3.4. As in the previous case, the critical values obtained vary depending on the magnitude of the estimate of the autoregressive coefficient in the data generating process.

Table 4.3.4. Empirical Critical Values and AR(1) Coefficients for Selected Commodity Price Series

Commodity	CVs / Trend	$\hat{\phi}$
Rubber	2.603	0.796
Timber	2,378	0.680
Lead	2.574	0.795

CV: Critical Value, $\hat{\phi}$: estimated autoregressive coefficient. The reported critical values are absolute values.

The estimates for the trend coefficient and the pertinent t-ratios are of course the same as those given in table 3.2.1. of chapter 3. These estimates together with the estimates for the absolute critical values obtained from the simulations above are reproduced in table 4.3.5.

Table 4.3.5 Estimated Trend Coefficients, t-ratios and Empirical 5% Critical Values for Selected Primary Commodity Price Series.

Commodity	Trend ($\hat{\beta}$)	t-ratio for $\hat{\beta}$	Critical Value
Beef	0.014	2.295	8.319
Cotton	-0.010	-3.328	8. 319
Lead	-0.006	-2.166	2.574
Maize	-0.010	-4.180	8. 319
Rubber	-0.028	-6.764	2.603
Timber	0.011	7.321	2.378
Wool	-0.016	-4.652	8. 319

$\hat{\beta}$: Estimated trend coefficient. The reported critical values are absolute values.

Of these estimates, the trend coefficients for Rubber and Timber appear to be statistically significant, although this would not have been so had the price series been modelled as stationary in first differences. All other price series do not appear to have a trend term significant at their relevant critical values. This is true even for the price series for Lead when the effect of serial correlation in moderately sized samples is taken into account although the coefficient would appear significant at the asymptotic normal critical value of +/-1.96.

4.3.3. Generalised Least Squares Estimators for the Trend Coefficient

An additional estimation method proposed by Sun and Pantula is a Generalised Least Squares regression of the form:

$$[4.3.4.] \qquad \begin{bmatrix} \sqrt{\left(1-\hat{\phi}^2\right)}\,Y_1 \\ Y_2-\hat{\phi}Y_1 \\ \vdots \\ Y_n-\hat{\phi}Y_{n-1} \end{bmatrix} = \begin{bmatrix} \sqrt{\left(1-\hat{\phi}^2\right)} & \sqrt{\left(1-\hat{\phi}^2\right)} \\ \left(1-\hat{\phi}\right) & \left[2-\hat{\phi}(2-1)\right] \\ \vdots & \vdots \\ \left(1-\hat{\phi}\right) & \left[n-\hat{\phi}(n-1)\right] \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

where, in the present case, both OLS and Maximum Likelihood estimates of ϕ have been used. It is obvious from equation [4.3.4.] that the EGLS (Estimated Generalised Least Squares) model reduces to a simple random walk with drift when $\hat{\phi} = 1$. Setting $\phi = 1$ in the data generating process, the empirical 5% critical value is given by +/-8,408 when ϕ is estimated through maximum likelihood and simulations with 10,000 replications are conducted for 99 observations under the null hypothesis of a zero trend coefficient[12]. This is reasonably close to Sun and Pantula's reported value of +/-8.42 for 100 observations. Using OLS to estimate the autoregressive parameter, the empirical upper and lower critical value is +/-8.891 compared to the critical value of +/-8.93 reported by Sun and Pantula.

As in the previous cases, the presence of large, positive autoregressive coefficients in the generating process results in non asymptotic critical values for the t-test statistic on the trend coefficient, for the given sample size. Table 4.3.6. below details the observed critical values obtained from simulations with 10,000

[12] t-ratios were obtained for the trend coefficient estimate obtained from fitting the EGLS model as in [4.3.4] and critical values were again inferred from the upper and lower 2.5 percentiles of the ranked t-ratios obtained by simulation.

replications each, for those commodities identified as trend stationary by either the Leybourne-McCabe test or the Augmented Dickey-Fuller test.

Table 4.3.6. Empirical Critical Values and AR(1) Coefficients for Selected Commodity Price Series

Commodity	CVs / OLS	$\hat{\phi}$ OLS	CVs / ML	$\hat{\phi}$ ML
Rubber	2,679	0.803	2,621	0.796
Timber	2,363	0.685	2,369	0.680
Lead	2,624	0.799	2,602	0.795

CV: Critical Value, $\hat{\phi}$: estimated autoregressive coefficient. The reported critical values are absolute values.

Results based on simulations where the autoregressive coefficient has been estimated by OLS are given in columns two and three. The corresponding results for the case where the autoregressive coefficient has been obtained by Maximum Likelihood estimation are presented in columns four and five of table 4.3.6. above.

It is not surprising that in those cases where the value of ϕ in the data generating process is one, and where consequently the EGLS model reduces to a random walk, the observed critical values are close to those obtained from previous simulations for the random walk case.

The 95% critical values for the adjusted data series with $\phi<1$ are also still above the asymptotic normal critical value of ±1.96, in fact, these values are also close to those obtained above by simple OLS or Maximum Likelihood estimation respectively.

Table 4.3.7. below reports the results of fitting the EGLS model to the original data series, estimating the autoregressive coefficient by OLS.

Table 4.3.7. Estimated Trend Coefficients, t-ratios and Empirical 5% Critical Values for Selected Primary Commodity Price Series.

Commodity	Trend ($\hat{b}$)	t-ratio for $\hat{b}$	Critical Value
Beef	0.013	2.137	8.891
Cotton	-0.010	-3.787	8.891
Lead	-0.006	-2.123	2,647
Maize	-0.010	-4.563	8.891
Rubber	-0.029	-6.478	2,601
Timber	0.011	7.084	2,374
Wool	-0.016	-5.352	8.891

$\hat{b}$: Estimated Trend Coefficient, Data series are adjusted for first order serial correlation
Estimation Method for ϕ: OLS. The reported critical values are absolute values.

The corresponding results for the case where the autoregressive coefficient is estimated by Maximum Likelihood estimation are given below in table 4.3.8:

Table 4.3.8. Estimated Trend Coefficients, t-ratios and Empirical 5% Critical Values for Selected Primary Commodity Price Series.

Commodity	Trend ($\hat{b}$)	t-ratio for $\hat{b}$	Critical Value
Beef	0.014	2.347	8.408
Cotton	-0.010	-3.341	8.408
Lead	-0.006	-2.152	2,619
Maize	-0.010	-4.066	8.408
Rubber	-0.029	-6,680	2,575
Timber	0.011	7.204	2,348
Wool	-0.016	-4.817	8.408

$\hat{b}$: Estimated Trend Coefficient, Data series are adjusted for first order serial correlation
Estimation Method for ϕ: Maximum Likelihood. The reported critical values are absolute values.

Again, the estimated trend coefficients for Rubber and Timber appear to be significant while those for Lead and the remaining difference stationary commodities do not. The t-ratio on the trend coefficient for Lead is close to the critical value obtained previously for moderately sized samples and above the asymptotic standard normal critical value of +/-1.96.

It seems then, that the finite sample correction of critical values for t-test statistics is not only a frequent problem in the case of serially correlated residuals, it also appears that the magnitude of this problem is difficult to assess. This in turn is related to the persistent underestimation of the AR(1) coefficient in the data generating process when fitting trend stationary models to a data series.

4.3.4. Adjusting for estimation bias in first order autoregression

To account for the observed downward bias in the estimation of $\hat{\phi}$, Sun and Pantula (*op. cit.*) propose a bias adjustment for the estimated coefficient, which is implemented as:

[4.3.5.] $$\tilde{\phi}_i = \hat{\phi}_i + \frac{2}{n} + \frac{2\hat{\phi}_i}{n}$$

The Generalised Least Squares estimates described above are then re-estimated using the bias adjusted autoregressive coefficient $\tilde{\phi}$. (Sun and Pantula suggest that in those cases where $\tilde{\phi} \geq 1$, the autoregressive coefficient in the EGLS equation should be set to one. In the present case however, this has not been relevant when the bias adjusted model was fitted to the original data series, since no values for $\tilde{\phi}$ in excess of one were obtained.) The subscript i refers to estimation by either OLS or Maximum Likelihood. The results obtained are summarised in table 4.3.9. below. Columns two and four list the estimated autoregressive coefficients for OLS and Maximum Likelihood estimation respectively, while the corresponding bias adjusted values are reported in columns three and five.

Table 4.3.9. Estimated and bias adjusted estimates for the autoregressive coefficient, using OLS and Maximum Likelihood estimation.

Commodity	$\hat{\phi}_{OLS}$	$\tilde{\phi}_{OLS}$	$\hat{\phi}_{ML}$	$\tilde{\phi}_{ML}$
Beef	0.914	0.953	0.905	0.944
Cotton	0.807	0.844	0.833	0.870
Lead	0.799	0.835	0.795	0.831
Maize	0.685	0.719	0.720	0.754
Rubber	0.803	0.840	0.796	0.832
Timber	0.685	0.719	0.680	0.714
Wool	0.801	0.838	0.824	0.861

$\hat{\phi}_{OLS}$: Estimate for the autoregressive coefficient, obtained by OLS. $\tilde{\phi}_{OLS}$ bias adjusted estimate of $\hat{\phi}_{OLS}$. $\hat{\phi}_{ML}$: Estimate for the autoregressive coefficient, obtained by ML. $\tilde{\phi}_{ML}$ bias adjusted estimate of $\hat{\phi}_{ML}$. ML: Maximum Likelihood.

Table 4.3.10. below reports the estimated trend coefficients and t-ratios obtained using EGLS with the bias adjusted estimate for the autoregressive coefficient. Columns two and three give the trend coefficient and t-ratio respectively for the OLS results, while the results from maximum likelihood estimation are reported in columns four and five.

Table 4.3.10. Estimated trend coefficients and t-ratios (bias adjusted method).

Commodity	$\hat{b}_{OLS}$	t-ratio	$\hat{b}_{ML}$	t-ratio
Beef	0.012	1.254	0.012	1.461
Cotton	-0.010	-3.143	-0.010	-2.681
Lead	-0.006	-1.820	-0.006	-1.850
Maize	-0.010	-4.082	-0.010	-3.566
Rubber	-0.029	-5.469	-0.029	-5.680
Timber	0.011	6.342	0.011	6.466
Wool	-0.016	-4.493	-0.016	-3.938

$\hat{b}$: estimated trend coefficient. ML: Maximum Likelihood

For the difference stationary time series, the critical value of +/-8.408 reported above can be employed here. For the trend stationary series, the implied empirical critical values were obtained by simulation. To conduct the simulations, data series where obtained constructing first order autoregressive residual series on the basis of normally distributed random numbers and using the adjusted value $\tilde{\phi}$ described above. Subsequently trend stationary models were fitted to the simulated series and the adjusted value $\tilde{\phi}$ was obtained as before. Again, $\tilde{\phi}$ was set to one in each replication where the value obtained exceeded one. Following this, the value of $\tilde{\phi}$ used in each replication was used to fit an EGLS model (according to [4.3.4.]) and retaining the t-ratio for the trend coefficient estimate obtained. Critical values were obtained as the top and bottom 2.5 percentiles of the t-ratios obtained by simulation and ranked in ascending order. The critical values obtained were +/-2.437 for Lead, +/-2.445 for Rubber and +/-2.267 for Timber when estimating ϕ by OLS. The corresponding results when $\hat{\phi}$ is obtained by Maximum Likelihood estimation are +/-2.415 for Lead, +/-2.413 for Rubber and +/-2.191 for Timber.

Among the difference stationary series there are again no statistically significant trend coefficients. The trend coefficients for Rubber and Timber do again appear to be significant, at the asymptotic critical value of +/-1.96 as well as at the critical values obtained by simulation, the estimated trend coefficient for Lead does not appear significant at either the asymptotic normal critical value, or at the empirical critical value of +/-2.425 obtained by simulation.

4.3.5. An alternative approach to *a priori* testing for the order of integration

In addition to the EGLS model shown in [4.3.4] and the bias adjusted estimate for the autoregressive coefficient according to [4.3.5] Sun and Pantula introduce a modified approach to pre-testing for unit roots.

Using the bias adjusted estimates for the first order autoregressive coefficient and the τ values obtained from augmented Dickey-Fuller tests, Sun and Pantula define a decision rule on the basis of which one may either assume that the data series in question is integrated (*i.e.* I(1)) or, alternatively, that the bias adjusted estimate of the AR(1) coefficient may be employed with or without further corrections, depending on the value of τ obtained in the unit root test.

More precisely, it is proposed that $\tilde{\phi}$ be adjusted according to:

[4.3.6.] $$\tilde{\phi}_i^* = C + (1 - C)\tilde{\phi}_i,$$

where the subscript i indicates estimation of $\tilde{\phi}$ by either OLS or Maximum Likelihood and C is defined as in:

[4.3.7.] $$C = \left\{ \begin{array}{ll} 0 & \text{if } \tau_i \leq -4.5, \\ 4.5 + \tau_i & -4.5 < \tau_i < -3.5, \\ 1 & \tau_i \geq -3.5 \end{array} \right|$$

with τ_i the Dickey Fuller test statistic obtained using either OLS or Maximum Likelihood estimates. It is obvious that for C=1 the autoregressive coefficient from [4.3.6.] will take a value of 1 also, while for C=0 the original bias adjusted estimate is retained.

Where the estimated autoregressive coefficient had to be adjusted further, the pertinent critical value was again obtained by simulation for 99 observations using 10,000 replications[13]. Critical values were again obtained as the top and bottom 2.5 percentiles of the t-ratios obtained by simulation and ranked in ascending order.

[13] The simulation methodology followed was mainly as in 4.3.4. above although on this occasion $\tilde{\phi}^*$, the adjusted version of the bias corrected AR(1) coefficient estimate was used in the data generating process.

The ADF test statistics obtained using OLS have been taken from the results presented in Appendix III.i. and are shown in table 4.3.11. below together with $\hat{\phi}$ obtained by OLS, the corresponding bias adjusted values and the final corrected values of the estimated autoregressive coefficient.

Table 4.3.11. Adjusted values for the OLS estimate of the autoregressive coefficient

Commodity	$\hat{\phi}_{OLS}$	$\tilde{\phi}_{OLS}$	τ_{OLS}	$\tilde{\phi}^{*}_{OLS}$
Beef	0.914	0.953	-1.876	1
Cotton	0.807	0.844	-2.553	1
Lead	0.799	0.835	-3.169	1
Maize	0.685	0.719	-2.468	1
Rubber	0.803	0.840	-3.217	1
Timber	0.685	0.719	-3.925	0.881
Wool	0.801	0.838	-2.137	1

$\hat{\phi}_{OLS}$:OLS estimate of the AR(1) coefficient, $\tilde{\phi}_{OLS}$: Adjusted coefficient value, τ_{OLS} : ADF test statistic $\tilde{\phi}^{*}_{OLS}$:Coefficient adjusted according to τ_{OLS}

It can be seen from table 4.3.11, that a difference stationary process would be inferred for all price series other than Timber, for which case the coefficient value would be adjusted further. This result is in accordance with the stationarity conclusions resulting from the Dickey Fuller test results reported in Chapter 3. The EGLS model represented by equation [4.3.4.] collapses to a random walk if the autoregressive coefficient is defined to be one and the corresponding OLS estimation results for all the AR(1) commodity price series other than Timber are listed in table 4.3.12. To obtain these results the data series in first differences were regressed on a constant.

Table 4.3.12. OLS and ML results for the EGLS model with ϕ=1

Commodity	$\hat{b}$	t-ratio	Commodity	$\hat{b}$	t-ratio
Beef	0.008	0.387	Maize	-0.008	-0.338
Cotton	-0.009	-0.510	Rubber	-0.030	-1.041
Lead	-0.008	-0.414	Wool	-0.014	-0.711

$\hat{b}$: estimated trend coefficient. The results for OLS and Maximum Likelihood are identical.

It is evident from table 4.3.12. that none of the estimated drift coefficients would be regarded as significant at the conventional critical value of +/-1.96 which should be appropriate in this case. For Timber the estimated trend coefficient was 0.881 with a t-ratio of 2.785. This would suggest the presence of a statistically significant trend term at the conventional critical value of +/-1.96 as well as at the critical value at the empirically determined critical value of 2.635, which had again been obtained by simulation using the same methodology as above (again for 99 observations over 10,000 replications).

Table 4.3.13. finally lists the original estimates of ϕ, the adjusted estimates and the ADF test statistics obtained using Maximum Likelihood estimation. (Details about ADF test results obtained using maximum likelihood estimation are given in appendix IV.iii.)

Table 4.3.13. Adjusted values for the Maximum Likelihood estimate of the autoregressive coefficient

Commodity	$\hat{\phi}_{ML}$	$\tilde{\phi}_{ML}$	τ_{ML}	$\tilde{\phi}^*_{ML}$
Beef	0.905	0.944	-1.876	1
Cotton	0.833	0.870	-1.835	1
Lead	0.795	0.831	-2.105	1
Maize	0.720	0.754	-2.210	1
Rubber	0.796	0.832	-3.104	1
Timber	0.680	0.714	-3.550	0.986
Wool	0.824	0.861	-2.137	1

$\hat{\phi}_{ML}$:Maximum Likelihood estimate of the AR(1) coefficient, $\tilde{\phi}_{ML}$: Adjusted coefficient value, τ_{ML} : ADF test statistic $\tilde{\phi}^*_{ML}$:Coefficient adjusted according toτ_{ML}

Again, all commodities other than Timber are classified as difference stationary. Regressing the first differenced commodity price series on a constant yields estimates identical to the ones reported in table 4.3.12. Estimating the EGLS equation for Timber using Maximum Likelihood estimation for ϕ yields an estimated trend coefficient value of 0.008 and a t-ratio of 0.702. The trend coefficient for Timber does not appear to be significant from this estimate at either the asymptotic critical value or the empirically determined value of +/-4.901. Furthermore, the estimated coefficient value is in accordance with the value of the drift coefficient reported in table 3.1 in chapter 3. The t-ratio is somewhat higher than the one in the difference stationary model selected by SBC. In contrast to the equivalent EGLS model estimated by OLS the adjusted value of the autoregressive coefficient is now closer to one than in the case of the OLS estimate reported above.

The methods used by Sun and Pantula (1999) allow for a more gradual approach in inferring the order of integration of a series, and in conjunction with simulation experiments allow one to give detailed consideration to the finite sample impact of serial correlation in the residual process. However, the method as considered here and in Sun and Pantula (*op. cit.*) is limited to rather simple model parameterisations, and does not resolve the problems associated with *a priori* assumptions on the order of integration in general.

4.4. An Alternative Testing Procedure for the Significance of a Trend Coefficient

In view of the above problems regarding significance tests for trend coefficients a testing procedure that is not sensitive to the presence of serial correlation or specific assumptions about the order of integration of the data generating process would greatly facilitate decisions on the presence of a trend term. Vogelsang (1998) proposes a test which reportedly is insensitive to serial correlation and adequate for data series which are either I(0) or I(1) even if the exact order of integration is not known.

4.4.1. Vogelsang's test for a trend coefficient

The starting point for Vogelsang's test is the simple trend stationary model:

[4.4.1] $$y_t = a + \beta t + u_t$$

where estimates of the coefficients α and β can be obtained by regressing y_t on a linear trend and constant. One should also note that it is not assumed here that the error term u_t is white noise. As a second step for the test, the partial sums $z_t = \Sigma_{j=1}^{T} y_t$ and $S_t = \Sigma_{j=1}^{T} u_t$ form part of the model:

[4.4.2] $$z_t = at + \beta^* \left[\tfrac{1}{2}(t^2 + t) \right] + S_t$$

which again can be estimated by OLS. In the following, the coefficient vectors for [4.4.1.] and [4.4.2.] will be referred to as $\widehat{\beta}$ and β^* respectively, likewise, the regressor matrices for [4.4.1.] and [4.4.2.] are referred to as **X1** and **X2** respectively. Also, extensions to models [4.4.1.] and [4.4.2.] are calculated in such a way as to add higher order trend polynomials to the basic OLS equations specified above. The resulting models are:

$$[4.4.3.] \qquad y_t = a + \hat{\beta} t + \sum_{j=2}^{m} \gamma_i t^i + u_t$$

and, for the regression of partial sums, [4.26]:

$$[4.4.4.] \qquad z_t = at + \beta^* \left[\tfrac{1}{2}(t^2 + t) \right] + \sum_{i=3}^{m} \gamma_i t^i + S_t$$

Analogous to the case for equations [4.4.1.] and [4.4.2.] $\tilde{\beta}_1$ and $\tilde{\beta}_2$ are used to refer to the vectors of coefficient estimates from [4.4.3.] and [4.4.4.] respectively[14]. One can then compute the OLS Wald statistics for [4.4.3.] and [4.4.4.] to test the hypothesis that $\gamma_2 = \gamma_3 = \ldots = \gamma_m = 0$ for [4.4.3.] and $\gamma_3 = \gamma_4 = \ldots = \gamma_m = 0$ for [4.4.4.]. Formulating the restriction as $\boldsymbol{Ri}\tilde{\beta}_i = r$, one would then need to specify

$$[4.4.5.] \qquad \boldsymbol{Ri} = \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

where the required dimensions of matrix $\boldsymbol{Ri}$ would be $\underset{m-1 \times m+1}{\boldsymbol{R1}}$ for [4.4.3.] and $\underset{m-2 \times m}{\boldsymbol{R2}}$ for [4.4.4.]. Defining the regressor matrix for [4.4.3.] as **X1a** and the regressor matrix for [4.4.4.] as **X2a** the OLS Wald statistic then takes the form:

$$[4.4.6.] \qquad \text{Wald}(i) = \frac{\left(\boldsymbol{Ri}\tilde{\beta}_i - r\right)' \left[\boldsymbol{Ri}(\boldsymbol{Xia}'\boldsymbol{Xia})^{-1} \boldsymbol{Ri}' \right]^{-1} \left(\boldsymbol{Ri}\tilde{\beta}_i - r\right)}{\tilde{\sigma}_j^2}$$

[14] *i.e.* the vector of coefficient estimates for the $\hat{\beta}$ and β^* coefficients as well as the estimated γ_i coefficients in equations [4.4.3.] and [4.4.4.].

with i=1,2, j=y_t, z_t and the restriction 'r' is defined to be r=0 in the present study. Normalising the Wald statistic over the number of observations (*i.e.* multiplying by T^{-1}, where in the present case T=99) yields the adjusted Wald statistic $J_T^i(m)$. Vogelsang (1998) then proposes two types of test statistic, the first of which can be written as:

$$[4.4.7.] \qquad PS_T^i = \frac{T^{-1}(c\beta^* - r)'\left[c(X_2'X_2)^{-1}c'\right]^{-1}(c\beta^* - r)}{S_z^2 \exp(bJ_T^i(m))}$$

where r is here set to r=0 as above and -since the interest is in testing the significance of the trend coefficient- c is set to take the form c=$\begin{bmatrix} 0 & 1 \end{bmatrix}$. Equation [4.4.7.] then reduces to:

$$[4.4.8.] \qquad PS_T^i = \frac{T^{-1}d_2^{-1}\beta^{*2}}{S_z^2 \exp(bJ_T^i(m))}, \; i = 1,2$$

here $d_2^{-1}$is the (2,2) element of the 2×2 matrix $(X_2'X_2)^{-1}$ and b is a constant. Both variants of the test use regressor matrix X_2 but PS_T^1 includes $J_T^1(m)$, while PS_T^2 makes use of $J_T^2(m)$. Vogelsang (1998) states that the asymptotic distributional characteristics of the normalised Wald test statistic $J_T^i(m)$ can be employed to make the test statistic specified in [4.4.8.] yield asymptotically equivalent results for series which are either I(0) or I(1), if the value of the constant *b* is specified appropriately[15]. Vogelsang (*op cit.*) provides values for b for different significance levels obtained by simulation. In the present study, the values for b chosen were b=1.966 for PS_T^1, and b=0.286 for PS_T^2. These values are appropriate for a test with

[15]More precisely, the distribution of the PS_T^i test statistic becomes right skewed for I(1) data series. This can be counter balanced as $\exp(bJ_T^i(m))$ takes on suitably large values (*cf.* Vogelsang (1998) pp. 130-131).

95% confidence in either case. The value of m was set at m=9 throughout, since the test is shown to have high power at this value (*cf.* Vogelsang (*op cit.*)).

The second type of test statistic takes the general form:

$$\text{[4.4.9.]} \qquad \mathrm{PSW}_T^i = \frac{(c\hat{\beta}-r)'\left[c(X_1'X_1)^{-1}c'\right]^{-1}(c\hat{\beta}-r)}{[T^{-1}100s_z^2 \exp(bJ_T^i(m))]}, \; i=1,2$$

where, as before, $c=\begin{bmatrix} 0 & 1 \end{bmatrix}$ and r=0. So that for the present study, [4.33] reduces to:

$$\text{[4.4.10.]} \qquad \mathrm{PSW}_T^i = \frac{d_1^{-1}\hat{\beta}^2}{[T^{-1}100s_z^2 \exp(bJ_T^i(m))]}, \; i=1,2$$

where now $d_1^{-1}$is the (2,2) element in the 2×2 matrix $(X_1'X_1)^{-1}$. For PSW_T^1, the normalised Wald statistic $J_T^1(m)$ has been used with m=9 and b=2.085 for a test with 95% confidence. Correspondingly, $J_T^2(m)$ has been used for PSW_T^2, where m=9 as well and the constant b now takes a value of b=0.305.

In order to assess the performance of the Vogelsang test, simulations have been conducted for a number of different AR(1) coefficient and trend values in the data generating process in a manner akin to the simulation experiment undertaken for ARIMA(1,0,0), ARIMA(0,1,0) and ARIMA (1,1,1) models in section 4.1 (*cf.* table 4.1.6.). The simulation results for the Vogelsang Test, again using 5000 replications, are reported below in table 4.4.1.

Table 4.4.1. Rejection Rates for Vogelsang Test Statistics, evidence from simulations with 5000 replications

ϕ	β	PS1	PS2	PSW1	PSW2
1	**0.00**	0.048	0.046	0.047	0.047
	0.05	0.058	0.063	0.058	0.060
	0.10	0.078	0.074	0.077	0.072
	0.20	0.130	0.111	0.128	0.111
0.95	***0.00***	0.025	0.034	0.025	0.032
	0.05	0.050	0.061	0.050	0.062
	0.10	0.107	0.105	0.108	0.104
	0.20	0.226	0.184	0.224	0.181
0.90	**0.00**	0.021	0.036	0.021	0.034
	0.05	0.093	0.103	0.096	0.104
	0.10	0.234	0.211	0.238	0.212
	0.20	0.441	0.328	0.438	0.323
0.80	***0.00***	0.026	0.041	0.025	0.042
	0.05	0.327	0.301	0.346	0.312
	0.10	0.661	0.500	0.674	0.498
	0.20	0.874	0.642	0.869	0.635
0.70	**0.00**	0.027	0.044	0.028	0.042
	0.05	0.660	0.529	0.682	0.535
	0.10	0.926	0.724	0.930	0.721
	0.20	0.989	0.834	0.990	0.827
0.60	***0.00***	0.028	0.042	0.029	0.042
	0.05	0.889	0.687	0.908	0.698
	0.10	0.991	0.847	0.992	0.844
	0.20	1.000	0.920	1.000	0.917

ϕ: value of the autoregressive coefficient in the data generating process, β: value of the trend coefficient in the data generating process, PS1 (4,537), PS2 (2,784), PSW1 (1,013) and PSW2 (0,612) are the Vogelsang test statistics (critical values are given in parentheses).

The simulation results show that, at least in the case of a first order autoregressive data generating process, the test has high power for large trend coefficients and does not suffer from the spurious rejection problem observed in ordinary trend stationary models for first order integrated and near integrated series. However the power of the Vogelsang test can be observed to fall substantially as the value of the trend coefficient decreases and for higher values of the autoregressive coefficient. For autoregressive coefficient values equal to or close to one, rejection rates for non-zero coefficients fall to very low levels. For the kind of low coefficient values obtained above for the GYCPI series the rejection rates for the I(1) case and for the case where ϕ=0.95 fall to values of around 5% and below. In view of the observed downward bias in estimating the autoregressive coefficient the possibility of low powered Vogelsang tests should therefore be borne in mind even in those cases where the estimated coefficients take values of around 0.9.

4.4.2. Results of the Vogelsang Test

Applying the Vogelsang test to the commodity price series underlying the GYCPI yielded the results given in table 4.4.2. Columns two to five list the values obtained for the test statistic identified in the header row. Those values which identify a trend coefficient as significant at the 5% level are identified by an asterisk. The 5% critical values differ by test statistic and are 4.537 for PS1, 2.784 for PS2, 1.013 for PSW1 and 0.612 for PSW2. The only two commodities for which a trend term is identified as unambiguously significant are Hides and Timber. (The series for Hides has been dropped in the remaining parts of this study since the series was no longer updated after 1995. Estimation results for ARIMA models for Hides are given in Appendix IV.iv.) There is, moreover, mixed evidence in favour of a trend for Lamb, where the PSW2 and PS2 statistics are significant, and for Aluminium, where the PSW1 and PS1 statistics suggest the presence of a significant trend term. For Wheat and Sugar, only the PSW1 statistic would suggest that the trend coefficient is significant.

Table 4.4.2 Test Results for the Vogelsang Test statistics

Commodity	PS1	PS2	PSW1	PSW2
Coffee	0.631	1.348	0.048	0.108
Cocoa	0.000	0.000	0.000	0.000
Tea	0.000	0.000	0.001	0.000
Rice	2.295	0.659	0.923	0.242
Wheat	3.423	0.349	1.238*	0.108
Maize	0.589	0.000	0.255	0.000
Sugar	3.391	1.271	1.031*	0.360
Beef	0.837	2.217	0.132	0.365
Lamb	3.087	9.886*	0.608	2.080*
Bananas	0.000	0.000	0.000	0.000
Palm oil	0.428	0.000	0.181	0.000
Cotton	0.215	0.005	0.089	0.001
Jute	0.000	0.000	0.004	0.000
Wool	0.079	0.000	0.025	0.000
Hides	8.168*	6.626*	2.920*	2.322*
Tobacco	0.001	0.000	0.000	0.00
Rubber	2.677	0.001	0.590	0.00
Timber	48.875*	48.287*	11.461*	11.265*
Copper	0.000	0.000	0.000	0.000
Aluminium	12.860*	1.845	2.700*	0.338
Tin	0.026	0.004	0.000	0.000
Silver	0.000	0.000	0.000	0.000
Lead	0.077	0.008	0.112	0.010
Zinc	0.191	0.053	0.013	0.003

Significant Values for the test statistic in question are identified by *. The 5% asymptotic critical values for the four test statistics in the table are: 4.537 for PS1, 2.784 for PS2, 1.013 for PSW1 and 0.612 for PSW2.

Bearing in mind that the trend coefficients obtained for the GYCPI series often take low values, a possible failure of the Vogelsang test to reject the null hypothesis of no trend for a non zero trend term should be considered. The results also don't provide evidence that would tempt one to classify one of the test statistics in question as inherently superior: for Aluminium, the null hypothesis is rejected by PS1 and PSW1 but not by PS2 or PSW2, while for Lamb the null hypothesis is rejected by PS2 and PSW2 but not by PS1 or PSW1.

4.4.3. Simulation evidence on the Vogelsang Test Statistics

To further assess the performance of the Vogelsang test for the commodities covered, simulation experiments were conducted for the commodity series in question[16]. Table 4.4.3. lists the rejection rates for the four Vogelsang test statistics, when data series where simulated for a model in levels without trend and the Vogelsang Test was applied subsequently to the simulated series over 5,000 replications.

[16] The simulation methodology used is again as described in 4.1. Coefficient estimates and residual standard errors used in the simulations were obtained from the ARIMA models selected by SBC as specified in column one of each table showing simulation results for the Vogelsang test.

Table 4.4.3. Rejection Rates of the Vogelsang Test Statistics When Series are Generated by a Trendless I(0) Process.

Commodity	ARIMA	PS1	PS2	PSW1	PSW2
Coffee	1,0,0	0.024	0.037	0.023	0.038
Cocoa	1,0,0	0.028	0.041	0.027	0.041
Tea	1,0,0	0.019	0.030	0.018	0.030
Rice	1,0,1	0.022	0.038	0.021	0.035
Wheat	0,0,3	0.024	0.045	0.024	0.047
Maize	1,0,0	0.027	0.041	0.026	0.043
Sugar	1,0,1	0.027	0.043	0.026	0.042
Beef	1,0,0	0.025	0.033	0.026	0.034
Lamb	5,0,0	0.005	0.018	0.004	0.019
Bananas	1,0,0	0.025	0.038	0.025	0.037
Palm oil	1,0,1	0.024	0.038	0.023	0.041
Cotton	1,0,0	0.027	0.036	0.027	0.036
Jute	1,0,0	0.024	0.037	0.022	0.036
Wool	1,0,0	0.024	0.035	0.023	0.032
Hides	1,0,0	0.020	0.038	0.019	0.035
Tobacco	1,0,0	0.030	0.039	0.030	0.038
Rubber	1,0,0	0.027	0.035	0.027	0.035
Timber	1,0,0	0.023	0.034	0.023	0.031
Copper	1,0,0	0.023	0.040	0.023	0.039
Aluminium	1,0,1	0.019	0.033	0.019	0.033
Tin	1,0,0	0.023	0.035	0.023	0.033
Silver	1,0,0	0.023	0.034	0.022	0.033
Lead	1,0,0	0.021	0.034	0.021	0.032
Zinc	1,0,1	0.028	0.045	0.029	0.043

The 5% asymptotic critical values for the four test statistics in the table are: 4.537 for PS1, 2.784 for PS2, 1.013 for PSW1 and 0.612 for PSW2. Simulations were conducted over 5000 replications.

Likewise, simulating from a difference stationary model without drift and subsequently computing the Vogelsang test over 5,000 replications yields the rejection rates given in table 4.4.4.

Table 4.4.4. Rejection Rates of the Vogelsang Test Statistics When Series are Generated by an I(1) Process without Drift.

Commodity	ARIMA	PS1	PS2	PSW1	PSW2
Coffee	0,1,0	0.046	0.050	0.045	0.047
Cocoa	2,1,0	0.077	0.055	0.075	0.054
Tea	0,1,0	0.055	0.055	0.055	0.055
Rice	1,1,2	0.017	0.012	0.017	0.012
Wheat	0,1,2	0.141	0.082	0.140	0.079
Maize	0,1,2	0.214	0.119	0.216	0.117
Sugar	0,1,2	0.191	0.102	0.189	0.099
Beef	0,1,0	0.050	0.054	0.049	0.052
Lamb	0,1,0	0.048	0.057	0.049	0.055
Bananas	0,1,0	0.053	0.053	0.053	0.053
Palm oil	2,1,2	0.097	0.064	0.097	0.062
Cotton	0,1,2	0.105	0.045	0.104	0.045
Jute	0,1,2	0.141	0.081	0.143	0.079
Wool	0,1,2	0.168	0.100	0.168	0.096
Hides	1,0,1	0.157	0.129	0.160	0.132
Tobacco	0,1,0	0.052	0.053	0.050	0.053
Rubber	0,1,0	0.051	0.054	0.050	0.055
Timber	0,1,0	0.055	0.055	0.053	0.055
Copper	0,1,0	0.051	0.057	0.050	0.055
Aluminium	1,1,2	0.129	0.104	0.127	0.103
Tin	0,1,0	0.051	0.055	0.051	0.053
Silver	2,1,0	0.085	0.061	0.085	0.060
Lead	0,1,0	0.046	0.049	0.045	0.047
Zinc	1,1,2	0.028	0.045	0.026	0.042

The 5% asymptotic critical values for the four test statistics in the table are: 4.537 for PS1, 2.784 for PS2, 1.013 for PSW1 and 0.612 for PSW2. Simulations were conducted over 5000 replications.

Tables 4.4.3. and 4.4.4. show that generally, the Vogelsang test tends to correctly fail to reject the null hypothesis of a zero trend coefficient when series are generated from a process without trend or drift coefficient regardless of whether the generating process is I(0) or I(1). Rejection rates for the case where a trendless series is generated by an I(1) process are somewhat higher than in the stationary case except in the case of Rice, where they are somewhat lower, and in the case of

Zinc where rejection rates are very close. It can also be observed that rejection rates are higher for I(1) models with an MA 2 component, where they take values around 10% or above. In most of the remaining cases, however, the rejection rates obtained remain reasonably close to a value of 5%.

It should be interesting then, to see how the test performs if data series are simulated including a trend using the coefficient and variance estimates obtained from fitting minimum SBC ARIMA models to the GYCPI data as above. The simulation results from doing so for trend stationary models with 5,000 replications are reported in table 4.4.5 below.

Table 4.4.5. Rejection Rates of the Vogelsang Test Statistics When Series are Generated by an I(0) Process with Trend.

Commodity	ARIMA	PS1	PS2	PSW1	PSW2
Coffee	1,0,0	0.060	0.078	0.064	0.082
Cocoa	1,0,0	0.028	0.047	0.029	0.045
Tea	1,0,0	0.311	0.251	0.311	0.249
Rice	1,0,1	0.715	0.612	0.739	0.614
Wheat	0,0,3	0.920	0.784	0.933	0.788
Maize	1,0,0	0.562	0.467	0.584	0.470
Sugar	1,0,1	0.706	0.598	0.741	0.621
Beef	1,0,0	0.130	0.139	0.132	0.140
Lamb	5,0,0	0.263	0.431	0.273	0.439
Bananas	1,0,0	0.033	0.044	0.033	0.043
Palm oil	1,0,1	0.626	0.546	0.655	0.556
Cotton	1,0,0	0.286	0.265	0.297	0.273
Jute	1,0,0	0.088	0.102	0.091	0.106
Wool	1,0,0	0.441	0.364	0.452	0.366
Hides	1,0,0	0.800	0.619	0.821	0.628
Tobacco	1,0,0	0.656	0.411	0.638	0.399
Rubber	1,0,0	0.675	0.508	0.682	0.506
Timber	1,0,0	0.883	0.685	0.889	0.685
Copper	1,0,0	0.056	0.072	0.054	0.072
Aluminium	1,0,1	0.889	0.716	0.895	0.715
Tin	1,0,0	0.023	0.037	0.022	0.036
Silver	1,0,0	0.025	0.037	0.023	0.037
Lead	1,0,0	0.169	0.183	0.179	0.188
Zinc	1,0,1	0.036	0.054	0.037	0.053

The 5% asymptotic critical values for the four test statistics in the table are: 4.537 for PS1, 2.784 for PS2, 1.013 for PSW1 and 0.612 for PSW2. Simulations were conducted over 5000 replications.

The corresponding results when simulating from an I(1) series with drift, again over 5000 replications, are given in table 4.4.6.

Table 4.4.6. Rejection Rates of the Vogelsang Test Statistics When Series are Generated by an I(1) Process with Drift.

Commodity	ARIMA	PS1	PS2	PSW1	PSW2
Coffee	0,1,0	0.053	0.058	0.056	0.056
Cocoa	2,1,0	0.095	0.068	0.096	0.069
Tea	0,1,0	0.057	0.059	0.059	0.058
Rice	1,1,2	0.599	0.452	0.612	0.456
Wheat	0,1,2	0.288	0.149	0.291	0.148
Maize	0,1,2	0.414	0.204	0.417	0.202
Sugar	0,1,2	0.262	0.131	0.262	0.131
Beef	0,1,0	0.052	0.050	0.052	0.050
Lamb	0,1,0	0.069	0.068	0.067	0.066
Bananas	0,1,0	0.052	0.056	0.051	0.056
Palm oil	2,1,2	0.103	0.069	0.102	0.069
Cotton	0,1,2	0.140	0.058	0.139	0.057
Jute	0,1,2	0.176	0.101	0.178	0.100
Wool	0,1,2	0.461	0.219	0.466	0.217
Hides	1,0,1	0.100	0.112	0.101	0.115
Tobacco	0,1,0	0.050	0.054	0.051	0.052
Rubber	0,1,0	0.075	0.078	0.075	0.076
Timber	0,1,0	0.055	0.057	0.054	0.056
Copper	0,1,0	0.056	0.063	0.056	0.061
Aluminium	1,1,2	0.853	0.704	0.858	0.702
Tin	0,1,0	0.052	0.052	0.051	0.050
Silver	2,1,0	0.085	0.062	0.083	0.058
Lead	0,1,0	0.058	0.056	0.059	0.055
Zinc	1,1,2	0.029	0.044	0.028	0.045

The 5% asymptotic critical values for the four test statistics in the table are: 4.537 for PS1, 2.784 for PS2, 1.013 for PSW1 and 0.612 for PSW2. Simulations were conducted over 5000 replications.

Rejection rates for a large number of trend stationary series are rather low often taking values equivalent to around 5% and surpassing 50% only for some commodities. Where the data generating process is difference stationary, higher rejection rates generally occur in models with an MA(2) component. There also appear to be differences in the rejection rates depending on the order of integration of the data generating process. Rejection rates for simulations with difference stationary data generating processes are generally somewhat lower than in the case of simulations where the DGP is I(0). Exceptions to this rule are the results for Cocoa, Bananas, Tin and Silver, where rejection rates are higher. Very close values for the observed rejection rates occur in the cases of Copper and Aluminium, for the PS2 and PSW2 values of Jute and the PS1 and PSW1 values for Wool.

Rejection probabilities, which are consistently above 0.5 are obtained for Rice and Aluminium. In the cases of Maize, Wheat, Sugar, Palm Oil, Hides, Rubber and Timber rejection probabilities are consistently above or at least close[17] to 0.5 when the DGP is stationary in levels and fall substantially below this value for difference stationary models. These observed differences in the rejection probabilities obtained seem to contradict Vogelsang's assertion that the test proposed is insensitive to the order of integration so long as the data generating process is either I(1) or I(0). At the very least, it appears that the test is sensitive to different model parameterisations rather than merely the significance of the trend component.

To further compare the simulation results for specific commodity series with the rejection rates given in table 4.4.1 the estimated trend or drift coefficient values can be divided by the standard error of the residual. This should facilitate comparison with the results in table 4.4.1 since for these a residual variance of 1 has been imposed on the data generating process. The adjusted trend coefficients

[17] Rejection probabilities somewhat below 0.5 are obtained for the PS2 and PSW2 statistics for Maize in the I(0) case with trend.

are given below in table 4.4.7. They were obtained using trend coefficient estimates and standard errors of the residuals from the minimum SBC models in levels.

Table 4.4.7. Normalised Values for Trend Coefficients

Commodity	Trend Coefficient	Commodity	Trend Coefficient
Coffee	0.018	Jute	-0.031
Cocoa	-0.012	Wool	-0.081
Tea†	-0.046	Hides*	-0.049
Rice	-0.069	Tobacco†	0.033
Wheat	-0.069	Rubber	-0.104
Maize	-0.049	Timber	0.076
Sugar	-0.034	Copper	-0.023
Beef	0.067	Aluminium	-0.124
Lamb	0.094	Tin	-0.005
Bananas	-0.012	Silver	-0.001
Palm Oil	-0.049	Lead	-0.033
Cotton	-0.061	Zinc	0.003

†Data up to 1997, *Data available up to 1995

Just over half of the adjusted coefficient estimates take absolute values corresponding to coefficients of around 0.05 or below in table 4.4.1. The remaining adjusted coefficient estimates fall mostly into the range of 0.5 to 0.1, although the normalised trend coefficient estimates for Aluminium and Rubber take values of -0.124 and -0.104 respectively. It can be concluded from the simulation results in table 4.4.1, that low rejection rates between 0.05-0.10 can be expected for first order integrated or near integrated series with coefficient values of this magnitude. However, for rigorous comparisons along these lines, one should concentrate on those cases for which the data generating process in the above simulation procedures is identified as ARIMA (1,0,0) by the Schwarz Bayesian Criterion. Table 4.4.8 below summarises the adjusted trend coefficient values as presented in table 4.4.7 above as well as the estimated values for the first order autoregressive coefficient from fitted ARIMA(1,0,0) models.

Table 4.4.8. Normalised Trend Coefficients and estimated AR(1) coefficient values for first order autoregressive price series.

Commodity	AR(1)	Trend Coefficient	Commodity	AR(1)	Trend Coefficient
Coffee	0.803	0.018	Tobacco†	0.953	0.033
Cocoa	0.870	-0.012	Hides*	0.633	-0.049
Tea†	0.883	-0.046	Rubber	0.796	-0.104
Maize	0.720	-0.049	Timber	0.680	0.076
Beef	0.905	0.067	Copper	0.856	-0.023
Bananas	0.926	-0.012	Tin	0.886	-0.005
Cotton	0.833	-0.061	Silver	0.884	-0.001
Jute	0.848	-0.031	Lead	0.795	-0.033
Wool	0.824	-0.081			

†Data up to 1997, *Data available up to 1995

Considering the estimates for the AR(1) coefficients, the simulation results obtained for the commodity price series covered seem to correspond closely to the results presented in table 4.4.1. Given the frequent combination of relatively low trend coefficient values and high positive autoregressive coefficients in the estimated trend stationary models, most series have rather low rejection rates for the Vogelsang test, ranging from values equivalent to around 0.3-0.4 in the case of Bananas to values of 0.8 and above for commodities such as Hides and Timber. In any cases rejection rates are well below 0.5 for most of the above series. Exceptions are Maize, where the estimated AR coefficient takes a value of 0.72 and Rubber, where the normalised trend coefficient with -0.100 takes a comparatively high absolute value. Indeed Hides and Timber, the two series for which the Vogelsang test unambiguously indicates the presence of a trend, and for which high rejection rates were obtained also, are characterised by their relatively low AR(1) coefficient values of 0.63 and 0.68 respectively, while the normalised trend coefficient estimates take values of 0.49 and 0.7 respectively[18].

This role of the AR(1) coefficient is crucial, in particular when comparing with the rejection rates for I(1) series. It can be seen immediately that the rejection rates even for the relatively large drift coefficient value of 0.2 scarcely surpass 0.10, (being close to 0.13 for the PS1 and PSW1 test statistics). For drift coefficient values of 0-0.05% rejection rates remain close to 0.05, rising to no more than *ca.* 0.07-0.08 for drift coefficient values around 0.1. The ratios of coefficient values and the standard error of the residual for the difference stationary models identified by minimum SBC are given in table 4.4.9. below.

Table 4.4.9. Adjusted values for drift coefficients

Commodity	Drift Coefficient	Commodity	Drift Coefficient
Coffee	0.009	Jute	-0.037
Cocoa	-0.038	Wool	-0.078
Tea†	-0.058	Hides*	-0.043
Rice	-0.072	Tobacco†	0.019
Wheat	-0.062	Rubber	-0.105
Maize	-0.048	Timber	0.050
Sugar	-0.038	Copper	-0.050
Beef	0.039	Aluminium	-0.126
Lamb	0.072	Tin	-0.017
Banana	0.004	Silver	-0.018
Palm Oil	-0.033	Lead	-0.042
Cotton	-0.053	Zinc	0.002

†Data up to 1997, *Data available up to 1995

Strict comparisons with the simulation results in table 4.16 are again only possible for those commodity series where the difference stationary model selected by minimum SBC is ARIMA (0,1,0). It is apparent, however, that almost all the

[18] It may be worth noting that the minimum SBC specifications for Hide identify a significant trend and drift term. The ARIMA(1,0,0) model specifies a trend coefficient of -0.012 with a t-ratio of -5.149, while the ARIMA(1,1,1) model -which is selected without imposing additional constraints- returns a drift coefficient estimate of -0.011 and a t-ratio of -2.827.

absolute coefficient values fall within the range 0-0.1. The exception is the normalised drift coefficient for Aluminium (-0.125) for which the I(1) ARIMA model is ARIMA (1,1,2). Nevertheless, in those cases where direct comparisons are possible, the low adjusted coefficient values lead one to expect low rejection rates within a 0.05-0.10 range or below. This is also confirmed by the rejection rates obtained for the simulations for the different commodity price series presented in table 4.4.6.

The implication of these simulation results is then that in those cases where there is a close correspondence between the rejection rates for the I(1) and I(0) models containing a trend term, the rejection rates for the series generated under the null and alternative hypotheses are rather close as well. This in turn implies that, under these circumstances, Vogelsang's test provides no basis on which to form predictions as to the presence of a significant trend term with a great degree of confidence, at least not in moderately and small sized samples. In those cases, where rejection rates for trend stationary models are high while those for I(1) models with drift are low, one is again faced with the necessity of deciding *a priori* on the order of integration. While spurious rejections are not a major problem when the Voglesang test is fitted to either stationary or difference stationary processes, low power in testing for significant drift terms remains an issue.

Given this difference in performance, one may then assess the Vogelsang test against the background of unit root pre-test results. Among the 24 commodities, those identified as trend stationary by ADF tests are Sugar, Lamb, Timber, Zinc, Aluminium and Hides. Among these, higher rejection rates, *i.e.* rejection rates noticeably above 50%, are obtained for Sugar, Timber, Aluminium and Hides when simulating from a stationary data generating process with trend. It is worth remembering then that the two commodities for which the Vogelsang Test unambiguously indicates the presence of a trend, Timber and Hides, have rejection rates between 0.6 and 0.9 when simulations are based on trend stationary series. The corresponding rejection rates for Lamb are 0.263 and 0.273 for the PS1 and PSW1 statistics respectively, while the rejection rates for the PS2 and PSW2 statistics are 0.431 and 0.439 respectively. It is, not surprisingly, on the basis of the PS2 and PSW2 statistic, that the Vogelsang Test tends to indicate the presence of a trend for Lamb. In the case of Aluminium, however, only the PS1 and PSW1 test statistics take values supportive of a trend, while the Vogelsang test statistics have rejection rates of 0.889 and 0.895 for the PS1 and PSW1 statistics respectively and values of 0.716 and 0.715 for the PS2 and PSW2 statistics (again, when the DGP is stationary in levels). For sugar, only the PSW2 statistic seems to indicate the possible presence of a trend, although the rejection rates obtained for simulations from a trend stationary model yield values of 0.706 and 0.741 for the PS1 and PSW1 statistics respectively and 0.598 and 0.621 for the PS2 and PSW2 statistics. For Zinc, finally, the Vogelsang test provides no support for the presence of a deterministic trend.

Considering further those price series identified as trend stationary by the Leybourne McCabe test, high rejection rates of 0.50 or above are obtained for Wheat, Rubber and Palm Oil when simulating from a trend stationary model. For Lead finally, rejection rates fall to within 0.16-0.19, while for Coffee they fall to values of around 0.06 to 0.08. Among these models, the only case where Vogelsang's Test gives at least some weak indication in favour of a significant trend is in the case of Wheat, where the PSW1 statistic is significant.

As pointed out above, rejection rates for difference stationary series with drift are below 50% in most cases and tend to fall below 0.10 for random walk plus drift models.

4.5 Conclusion

Giving further consideration to the issue of trend measurement in primary commodity price series and the order of integration of these series, different conclusions on stationarity can be arrived at if a stationarity test is applied. Using the Leybourne McCabe test the price series for Coffee, Wheat, Palm Oil, Rubber and Lead can now be classified as trend stationary while they would be identified as difference stationary on the basis of the ordinary ADF test. It should be noted though that some fundamental problems of uncertainty surrounding *a priori* testing in general still persist in this case. The Leybourne McCabe test, while accounting for large moving average roots, does not remove the fundamental problems of essentially arbitrarily specified critical values for any test statistic. The issue of possible structural instability and multiple structural breaks also remains unresolved.

Aside from the issues of uncertainty surrounding *a priori* testing, it was shown that spurious rejections of the null hypothesis of a zero trend coefficient are a problem in near integrated time series with first order autoregressive coefficients as low as 0.7. Some progress can be made in the case of AR(1) series, if one attempts to correct the critical values used in hypothesis tests for the effect of serial correlation, yet there remain problems due to the fact that estimates of the first order autoregressive coefficient tend to be biased.

In the case of difference stationary models a frequent problem is a loss of power of the conventional t-test if the model selected by SBC is underparameterised. It has been attempted to avoid this problem by either fitting an ARIMA(1,1,1) model instead of random walk plus drift as an alternative to a first order autoregressive trend stationary model, by constraining the minimum SBC values or selecting the appropriate model by AIC, with selection by AIC being the more generally applicable approach, since it has not been confined to comparing ARIMA (1,0,0) with ARIMA(0,1,0) models. This method somewhat increases the danger of spurious rejections of the null hypothesis for integrated series[19] and yields t-tests which are more powerful than t-tests for underparameterised difference stationary models. It also avoids the problem of spurious rejection for near integrated series, although the t-test on the drift coefficient is still less powerful than for the t-test in the trend stationary model selected by SBC.

An alternative testing procedure proposed by Vogelsang is said to be applicable for time series which are either trend stationary or I(1). Application of the test to the present data series and further simulation experiments do reveal however, that the results obtained are far from unambiguous. Vogelsang's test was shown to have low power in integrated and near integrated series and where the estimated coefficient values are low. The power of the test is high for AR(1) series with low

[19] This has at least been shown for the case of ARIMA (1,1,1) models.

values for the autoregressive coefficient, while avoiding spurious rejections of the null hypothesis when the trend coefficient is actually zero. This difference in performance does however imply that *a priori* knowledge of the order of integration and knowledge of the value of the autoregressive coefficient in the data generating process become important in assessing the reliability of the test. Hence, uncertainty about *a priori* testing for stationarity and the bias in measuring the magnitude of the autoregressive coefficient remain an issue with the Voglesang test. It has not been clear, moreover, which of the various test statistics proposed by Vogelsang is most appropriate, while the test results have not been uniform for all commodity price series.

The main conclusion from the tests conducted then seems to be that no one single testing procedure should be relied upon to yield conclusive results. Table 4.5.1. summarises the results from various tests on the trend coefficient.

Column two gives the drift coefficient from the difference stationary model selected by AIC. Conclusions on the significance of the trend coefficient from the t-ratio from the model in levels selected by SBC are given in column three. I(0) Models with a significant trend coefficient have been listed here only if either they were identified as trend stationary through *a priori* testing or if there was evidence in favour of a significant trend or drift coefficient from one of the other tests employed. Those series with no evidence in favour of a significant trend or drift coefficient have also been omitted from table 4.5.1. Column four summarises conclusions on the significance of the drift coefficient based on the t-ratio from the difference stationary model selected by AIC. The results of Vogelsang's test are summarised in column five, with the numbers in brackets indicating how many of the 4 tests indicated the presence of a significant trend. The column labelled S&P indicates if the various tests proposed by Sun and Pantula -other than the one listed in the final column- fail to reject the null hypothesis of a zero trend coefficient. In the final column labelled S&P2, the results of a test based on a modified pre-test method by Sun and Pantula (1999) are listed. This test does not reject the null hypothesis of a zero trend coefficient for any of the commodities covered, and yields inconclusive results for timber.

Table 4.5.1. Evidence in favour of a trend or drift coefficient from various tests.

Commodity	$\hat{\beta}$	t-ratio SBC/I(0)	t-ratio AIC/I(1)	Vogelsang	S&P	S&P2
Aluminium	-0.019	*	*	[2/4]	-	-
Hides	-0.011#	*	*	[4/4]	-	-
Lamb	0.015	*	/	[2/4]	-	-
Lead	-0.008	*	/	[0/4]	/	/
Palm Oil	-0.010	*	/	[0/4]	-	-
Rice	-0.012	*	*	[0/4]	-	-
Rubber	-0.029	*	*	[0/4]	*	/
Sugar	-0.011	*	*	[1/4]	-	-
Timber	0.008	*	*	[4/4]	*	?
Wheat	-0.011	*	(/)	[1/4]	-	-

#: Data for Hides end in 1995, "-": not applicable
*: Significant Trend / Drift coefficient, "/": Trend / Drift coefficient not significant,
(/): Trend / Drift coefficient marginally insignificant, ?: Test inconclusive

Following the results presented in the table, one may conclude with reasonable confidence, that a significant trend or drift term is present in the price series of Aluminium, Hides, Rubber, Sugar and Timber. Evidence in favour of a trend is strongest for Timber and Hides. There may be some doubt about this conclusion in the case of Rubber, since none of the Vogelsang test statistics rejects the null hypothesis, although some of the Sun and Pantula tests do.

The case is less clear for Rice, Lamb and Wheat. The Vogelsang test statistics give ambiguous results for Lamb and Wheat providing stronger evidence against a trend in the case of Wheat. The price series for Wheat is identified as difference stationary by the ADF test and as trend stationary by the Leybourne McCabe test. The t-ratio on the drift coefficient takes a value of -1.923 when the difference stationary model is selected by AIC, thus showing the trend coefficient to be only marginally insignificant. For Rice, rejections of the null hypothesis are obtained from the t-tests in the trend stationary model and from the difference stationary model when the latter is selected by either SBC or AIC. Sun and Pantula's tests are not applicable in this case, however, and none of the Vogelsang tests reject the null hypothesis. In the cases of Lead and Palm Oil one is tempted to accept the null hypothesis of a zero trend coefficient, since only one test -the t-ratio from the trend stationary model- identifies the trend coefficient as significant.

Adopting a cautious attitude, the presence of a trend or drift in the relative price series can only be inferred for seven of the twenty four commodities: Aluminium, Hides, Rice, Rubber, Sugar, Timber and Wheat. All of these price series also show some evidence in favour of a significant drift coefficient in the difference stationary model selected by AIC. No one particular test seems to be sufficient for conclusions about the significance of the trend coefficient, so that it appears most adequate to consider evidence from a variety of testing procedures when deciding on the significance of a trend term. When selecting models on the basis of information criteria while allowing for a trend or drift term, one should consider the AIC for the selection of difference stationary representations of the data generating process while focusing on the SBC in the trend stationary case.

Appendix IV.i. Estimation Output for ARIMA Models with Higher Parameterisations

In section 4.1.2. it was outlined why the selection of overly parsimonious model specifications by the SBC may lead to mistaken inferences about the presence of a drift term and two methods for selecting more elaborate presentations of the underlying time series were employed. This appendix provides the full estimated equations underlying the summary output in tables 4.1.7. and 4.1.8.

IV.i.i. ARIMA(p,1,q) models with p>=1 and q>=1, selected by SBC.

Table 4.1.7. summarised the estimation results for drift coefficients and t-ratios obtained when ARIMA models are selected by the Schwarz Bayaesian Criterion while the minimum number of autoregressive lags (p) and the minimum number of moving average terms (q) are constrained to be at least one. The full estimated equations for the relevant commodities, *i.e.* those for which an ARIMA(1,0,0) model was specified in levels while the same selection process for difference stationary models yielded ARIMA(0,1,0) specifications, are specified below. The standard errors are again indicated in parenthesis below the coefficient estimates. In each case Ljung Box Q statistic for 12 autocorrelations is reported with P-values given in parentheses[1].

[1] For the ARIMA(1,1,1) models for Lamb (and Tea with 6.6%) this indicates the possible presence of autocorrelated residuals.

Commodity: Bananas:
Model: ARIMA(1,1,1) with constant
SBC = -179.663 , **Ljung-Box Q(12)**:8.933 (0.538)
Degrees of freedom: 95

$$\Delta p_t = \underset{(0.010)}{0.000} + v_t$$

$$v_t + \underset{(0.661)}{0.430}\, v_{t-1} = \underset{(0.092)}{\varepsilon_t} + \underset{(0.618)}{0.538}\, \varepsilon_{t-1}$$

Commodity: Tobacco
Model: ARIMA(3,1,2) with constant
SBC = -92.650 , **Ljung-Box Q(12)**:6.308 (0.504)
Degrees of freedom: 91

$$\Delta p_t = \underset{(0.013)}{0.003} + v_t$$

$$v_t - \underset{(0.115)}{1.272}\, v_{t-1} + \underset{(0.168)}{0.823}\, v_{t-2} + \underset{(0.111)}{0.118}\, v_{t-3} = \underset{(0.131)}{\varepsilon_t} - \underset{(n.a.)}{1.349}\, \varepsilon_{t-1} + \underset{(n.a.)}{\varepsilon_{t-2}}$$

Commodity: Rubber:
Model: ARIMA(1,1,2) with constant
SBC = 41.130 , **Ljung-Box Q(12)**:5.783 (0.761)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.004)}{0.029} + v_t$$

$$v_t - \underset{(0.104)}{0.769}\, v_{t-1} = \underset{(0.274)}{\varepsilon_t} - \underset{(n.a.)}{0.827}\, \varepsilon_{t-1} - \underset{(n.a.)}{0.173}\, \varepsilon_{t-2}$$

Commodity: Timber:
Model: ARIMA(1,1,1) with constant
SBC = -75.277 , **Ljung-Box Q(12)**:9.000 (0.532)
Degrees of freedom: 95

$$\Delta p_t = \underset{(0.006)}{0.011} + v_t$$

$$v_t - \underset{(0.149)}{0.848}\, v_{t-1} = \underset{(0.154)}{\varepsilon_t} - \underset{(n.a.)}{\varepsilon_{t-1}}$$

Commodity: Copper
Model: ARIMA(1,1,2) with constant
SBC = -37.411 , **Ljung-Box Q(12)**:2.916 (0.968)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.011)}{0.008} + v_t$$

$$v_t - \underset{(0.288)}{0.484}\, v_{t-1} = \underset{(0.186)}{\varepsilon_t} - \underset{(0.287)}{0.467}\, \varepsilon_{t-1} - \underset{(0.113)}{0.237}\, \varepsilon_{t-2}$$

Commodity: Tin
Model: ARIMA(1,1,1) with constant
SBC = -34.893 , **Ljung-Box Q(12)**:5.020 (0.890)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.021)}{0.004} + v_t$$

$$v_t + \underset{(1.559)}{0.079}\, v_{t-1} = \underset{(0.192)}{\varepsilon_t} + \underset{(1.548)}{0.144}\, \varepsilon_{t-1}$$

Commodity: Lead
Model: ARIMA(1,1,1) with constant
SBC = -43.721 , **Ljung-Box Q(12)**:9.992 (0.441)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.005)}{0.005} + v_t$$

$$v_t - \underset{(0.107)}{0.851}\, v_{t-1} = \underset{(0.181)}{\varepsilon_t} - \underset{(n.a.)}{\varepsilon_{t-1}}$$

IV.i.ii. ARIMA(1,1,1) models

ARIMA(1,1,1) models are estimated for Tea, Tobacco and Copper, standard errors are given in parentheses below the coefficient estimates.

Commodity: Tea
Model: ARIMA(1,1,1) with constant
SBC = -63.689 , **Ljung-Box Q(12):**17.404 (0.066)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.017)}{0.010} + v_t$$

$$v_t + \underset{(0.169)}{0.982}\, v_{t-1} = \underset{(0.164)}{\varepsilon_t} - \underset{(n.a.)}{\varepsilon_{t-1}}$$

Commodity: Tobacco
Model: ARIMA(1,1,1) with constant
SBC = -91.787 , **Ljung-Box Q(12):**10.924 (0.364)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.015)}{0.003} + v_t$$

$$v_t - \underset{(1.830)}{0.160}\, v_{t-1} = \underset{(0.143)}{\varepsilon_t} - \underset{(1.184)}{0.103}\, \varepsilon_{t-1}$$

Commodity: Copper
Model: ARIMA(1,1,1) with constant
SBC = -36.597 , **Ljung-Box Q(12):**6.891 (0.736)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.021)}{0.009} + v_t$$

$$v_t + \underset{(0.744)}{0.400}\, v_{t-1} = \underset{(0.190)}{\varepsilon_t} + \underset{(0.704)}{0.504}\, \varepsilon_{t-1}$$

IV.i.iii. ARIMA(p,1,q) models selected by AIC

The difference stationary model specifications selected by AIC, which were computed for all commodities are as follows (standard errors are again in parentheses below the estimated coefficient values):

Commodity: Coffee
Model: ARIMA(0,1,2) with constant
AIC = 7.121 , **Ljung-Box Q(12):**6.668 (0.756)
Degrees of freedom: 95

$$\Delta p_t = \underset{(0.018)}{0.003} + v_t$$

$$v_t = \underset{(0.247)}{\varepsilon_t} - \underset{(0.100)}{0.043}\, \varepsilon_{t-1} - \underset{(0.102)}{0.255}\, \varepsilon_{t-2}$$

Commodity: Cocoa
Model: ARIMA(2,1,0) with constant
AIC = 5.097 , **Ljung-Box Q(12):**5.913 (0.823)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.020)}{0.009} + v_t$$

$$v_t - \underset{(0.097)}{0.082}\, v_{t-1} + \underset{(0.097)}{0.311}\, v_{t-2} = \underset{(0.244)}{\varepsilon_t}$$

Commodity: Tea
Model: ARIMA(0,1,2) with constant
AIC = -77.963 , **Ljung-Box Q(12):**9.869 (0.452)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.011)}{0.011} + v_t$$

$$v_t = \underset{(0.159)}{\varepsilon_t} - \underset{(0.099)}{0.044}\, \varepsilon_{t-1} - \underset{(0.099)}{0.300}\, \varepsilon_{t-2}$$

Commodity: Rice
Model: ARIMA(1,1,2) with constant
AIC = -71.385 , **Ljung-Box Q(12)**:8.037 (0.530)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.005)}{0.012} + v_t$$

$$v_t - \underset{(0.140)}{0.551}\, v_{t-1} = \underset{(0.164)}{\varepsilon_t} - \underset{(0.131)}{0.349}\, \varepsilon_{t-1} - \underset{(0.093)}{0.542}\, \varepsilon_{t-2}$$

Commodity: Wheat
Model: ARIMA(0,1,4) with constant
AIC = -84.132 , **Ljung-Box Q(12)**:6.880 (0.550)
Degrees of freedom: 93

$$\Delta p_t = -\underset{(0.006)}{0.011} + v_t$$

$$v_t = \underset{(0.153)}{\varepsilon_t} + \underset{(0.100)}{0.092}\, \varepsilon_{t-2} - \underset{(0.103)}{0.380}\, \varepsilon_{t-2} - \underset{(0.102)}{0.068}\, \varepsilon_{t-3} - \underset{(0.102)}{0.308}\, \varepsilon_{t-4}$$

Commodity: Maize
Model: ARIMA(0,1,2) with constant
AIC = -29.228 , **Ljung-Box Q(12)**:11.805 (0.298)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.007)}{0.10} + v_t$$

$$v_t = \underset{(0.205)}{\varepsilon_t} - \underset{(0.093)}{0.218}\, \varepsilon_{t-1} - \underset{(0.093)}{0.441}\, \varepsilon_{t-2}$$

Commodity: Sugar
Model: ARIMA(0,1,5) with constant
AIC = 55.357 , **Ljung-Box Q(12)**:8.316 (0.306)
Degrees of freedom: 92

$$\Delta p_t = -\underset{(0.003)}{0.011} + v_t$$

$$v_t = \underset{(0.306)}{\varepsilon_t} - \underset{(n.a.)}{0.064}\, \varepsilon_{t-1} - \underset{(n.a.)}{0.521}\, \varepsilon_{t-2} - \underset{(n.a.)}{0.141}\, \varepsilon_{t-3} + \underset{(n.a.)}{0.024}\, \varepsilon_{t-4} - \underset{(n.a.)}{0.298}\, \varepsilon_{t-5}$$

Commodity: Beef
Model: ARIMA(0,1,0) with constant
AIC = -30.383 , **Ljung-Box Q(12):**10.553 (0.568)
Degrees of freedom: 97

$$\Delta p_t = \underset{(0.021)}{0.008} + v_t$$

$$v_t = \underset{(0.206)}{\varepsilon_t}$$

Commodity: Lamb
Model: ARIMA(4,1,1) with constant
AIC = -26.276 , **Ljung-Box Q(12):**4.540 (0.716)
Degrees of freedom: 92

$$\Delta p_t = \underset{(0.026)}{0.015}\, v_t$$

$$v_t + \underset{(0.233)}{0.404}\, v_{t-1} + \underset{(0.107)}{0.040}\, v_{t-2} + \underset{(0.108)}{0.018}\, v_{t-3} - \underset{(0.106)}{0.321}\, v_{t-4} = \underset{(0.205)}{\varepsilon_t} + \underset{(0.241)}{0.455}\, \varepsilon_{t-1}$$

Commodity: Bananas
Model: ARIMA(0,1,0) with constant
AIC = -190.022 , **Ljung-Box Q(12):**10.778 (0.548)
Degrees of freedom: 97

$$\Delta p_t = \underset{(0.009)}{0.000} + v_t$$

$$v_t = \underset{(0.091)}{\varepsilon_t}$$

Commodity: Palm Oil
Model: ARIMA(0,1,3) with constant
AIC = -21.176 , **Ljung-Box Q(12):**7.842 (0.550)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.009)}{0.010}\, v_t$$

$$v_t = \underset{(0.212)}{\varepsilon_t} + \underset{(0.101)}{0.032}\, \varepsilon_{t-1} - \underset{(0.091)}{0.429}\, \varepsilon_{t-2} - \underset{(0.101)}{0.210}\, \varepsilon_{t-3}$$

Commodity: Cotton
Model: ARIMA(2,1,3) with constant
AIC = -89.016 , **Ljung-Box Q(12):**5.582 (0.589)
Degrees of freedom: 92

$$\Delta p_t = -\underset{(0.012)}{0.007} + v_t$$

$$v_t - \underset{(0.131)}{1.130}\, v_{t-1} + \underset{(0.130)}{0.595}\, v_{t-2} = \underset{(0.148)}{\varepsilon_t} - \underset{(0.152)}{1.323}\, \varepsilon_{t-1} + \underset{(0.243)}{0.408}\, \varepsilon_{t-2} + \underset{(0.130)}{0.285}\, \varepsilon_{t-3}$$

Commodity: Jute
Model: ARIMA(0,1,2) with constant
AIC = -20.423 , **Ljung-Box Q(12):**13.184 (0.214)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.012)}{0.008} + v_t$$

$$v_t = \underset{(0.214)}{\varepsilon_t} - \underset{(0.095)}{0.053}\, \varepsilon_{t-1} - \underset{(0.095)}{0.399}\, \varepsilon_{t-2}$$

Commodity: Wool
Model: ARIMA(0,1,2) with constant
AIC = -48.509 , **Ljung-Box Q(12):**3.771 (0.957)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.008)}{0.014} + v_t$$

$$v_t = \underset{(0.186)}{\varepsilon_t} - \underset{(0.094)}{0.172}\, \varepsilon_{t-1} - \underset{(0.094)}{0.420}\, \varepsilon_{t-2}$$

Commodity: Tobacco
Model: ARIMA(3,1,2) with constant
AIC = -108.098 , **Ljung-Box Q(12):**6.308 (0.504)
Degrees of freedom: 91

$$\Delta p_t = \underset{(0.013)}{0.003} + v_t$$

$$v_t - \underset{(0.115)}{1.272}\, v_{t-1} + \underset{(0.168)}{0.823}\, v_{t-2} + \underset{(0.111)}{0.118}\, v_{t-3} = \underset{(0.131)}{\varepsilon_t} - \underset{(n.a.)}{1.349}\, \varepsilon_{t-1} + \underset{(n.a.)}{\varepsilon_{t-2}}$$

Commodity: Rubber
Model: ARIMA(1,1,2) with constant
AIC = 30.790 , **Ljung-Box Q(12)**:5.783 (0.761)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.004)}{0.029} + v_t$$

$$v_t - \underset{(0.104)}{0.769}\, v_{t-1} = \underset{(0.274)}{\varepsilon_t} - \underset{(n.a.)}{0.827}\, \varepsilon_{t-1} - \underset{(n.a.)}{0.173}\, \varepsilon_{t-2}$$

Commodity: Timber
Model: ARIMA(0,1,5) with constant
AIC = -84.650 , **Ljung-Box Q(12)**:2.435 (0.932)
Degrees of freedom: 92

$$\Delta p_t = \underset{(0.001)}{0.012} + v_t$$

$$v_t = \underset{(0.150)}{\varepsilon_t} - \underset{(n.a.)}{0.220}\, \varepsilon_{t-1} - \underset{(n.a.)}{0.321}\, \varepsilon_{t-2} - \underset{(n.a.)}{0.034}\, \varepsilon_{t-3} - \underset{(n.a.)}{0.183}\, \varepsilon_{t-4} - \underset{(n.a.)}{0.241}\, \varepsilon_{t-5}$$

Commodity: Copper
Model: ARIMA(1,1,2) with constant
AIC = -47.751 , **Ljung-Box Q(12)**:2.916 (0.968)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.011)}{0.008} + v_t$$

$$v_t - \underset{(0.288)}{0.484}\, v_{t-1} = \underset{(0.186)}{\varepsilon_t} - \underset{(0.287)}{0.467}\, \varepsilon_{t-1} - \underset{(0.113)}{0.237}\, \varepsilon_{t-2}$$

Commodity: Aluminium
Model: ARIMA(1,1,2) with constant
AIC = -84.632 , **Ljung-Box Q(12)**:2.965 (0.966)
Degrees of freedom: 94

$$\Delta p_t = -\underset{(0.002)}{0.019} + v_t$$

$$v_t - \underset{(0.103)}{0.679}\, v_{t-1} = \underset{(0.152)}{\varepsilon_t} - \underset{(n.a.)}{0.537}\, \varepsilon_{t-1} - \underset{(n.a.)}{0.463}\, \varepsilon_{t-2}$$

Commodity: Tin
Model: ARIMA(0,1,0) with constant
AIC = -46.271 , **Ljung-Box Q(12):**5.395 (0.943)
Degrees of freedom: 97

$$\Delta p_t = -\underset{(0.019)}{0.003} + v_t$$

$$v_t = \underset{(0.190)}{\varepsilon_t}$$

Commodity: Silver
Model: ARIMA(2,1,0) with constant
AIC = -54.185 , **Ljung-Box Q(12):**8.089 (0.620)
Degrees of freedom: 95

$$\Delta p_t = -\underset{(0.014)}{0.003} + v_t$$

$$v_t - \underset{(0.098)}{0.041}\, v_{t-1} + \underset{(0.098)}{0.307}\, v_{t-2} = \underset{(0.181)}{\varepsilon_t}$$

Commodity: Lead
Model: ARIMA(0,1,4) with constant
AIC = -52.185 , **Ljung-Box Q(12):**4.736 (0.785)
Degrees of freedom: 93

$$\Delta p_t = -\underset{(0.007)}{0.008} + v_t$$

$$v_t = \underset{(0.180)}{\varepsilon_t} - \underset{(0.102)}{0.092}\, \varepsilon_{t-1} - \underset{(0.101)}{0.134}\, \varepsilon_{t-2} - \underset{(0.101)}{0.246}\, \varepsilon_{t-3} - \underset{(0.104)}{0.178}\, \varepsilon_{t-4}$$

Commodity: Zinc
Model: ARIMA(1,1,2) with constant
AIC = -32.835 , **Ljung-Box Q(12):**6.374 (0.702)
Degrees of freedom: 94

$$\Delta p_t = \underset{(0.002)}{0.000} + v_t$$

$$v_t - \underset{(0.138)}{0.478}\, v_{t-1} = \underset{(0.197)}{\varepsilon_t} - \underset{(n.a.)}{0.639}\, \varepsilon_{t-1} - \underset{(n.a.)}{0.361}\, \varepsilon_{t-2}$$

Appendix IV.ii. Further Details on the Leybourne-McCabe Stationarity Test

IV.ii.i. Critical Values for Stationarity Tests

Kwiatkowski *et.al.* (1992) provide a number of critical values for stationarity tests like those developed by Kwiatkowski *et.al.* (*op.cit.*) and Leybourne and McCabe (1994) or Leybourne and McCabe (1999). Table IV.ii.i. details the critical values for the test statistic for a stationarity test including a trend $(\tilde{S}_\beta(p))$.

Table IV.ii.i. Upper tail Percentiles and Critical Values for the $(\tilde{S}_\beta(p))$ Test Statistic

Signif. Level	0.100	0.050	0.025	0.010
Critical Value	0.119	0.146	0.176	0.216

Source: Kwiatkowski *et. al.* 1992 *Testing the Null Hypothesis of Trend Stationarity*, Journal of Econometrics, Vol. 54, pp159-178

IV.ii.ii. ARIMA(p,1,1) Equations used in the Leybourne McCabe Test.

Table 4.2.1 specifies the final values for the $\tilde{S}_\beta(p)$ statistic obtained. This appendix lists the full specifications for the autoregressive equations used in computing the Leybourne McCabe test statistic. (The autoregressive terms are calculated for the residual rather than in terms of lagged dependent variables. It was shown in Chapter 2 though that the autoregressive and moving average coefficients should be equivalent in both cases.) Along with the estimated equations, the Ljung Box Q statistic for eight autocorrelations is given with P-values in parentheses. (Q(8) was preferred over Q(6) since the latter does not allow one to compute the P-value for testing equations with five autocorrelations or more. For a similar reason, Q(10) is reported in the case of Rice.)[1] The Z statistic on the last autoregressive parameter (as in Leybourne and McCabe (1999)) is also reported where appropriate.

Coffee: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 6.204 (0.401), **Z**=9.266

$$\Delta p_t = \underset{(0.012)}{0.005} + v_t$$

$$v_t - \underset{(0.172)}{0.931} v_{t-1} = \underset{(0.247)}{\varepsilon_t} - \underset{(\ldots)}{\varepsilon_{t-1}}$$

Cocoa: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 10.191 (0.117), **Z**=9.152

$$\Delta p_t = -\underset{(0.026)}{0.008} + v_t$$

$$v_t + \underset{(0.086)}{0.920} = \underset{(0.252)}{\varepsilon_t} + \underset{(\ldots)}{\varepsilon_{t-1}}$$

Tea: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 10.500 (0.105), **Z**=9.726

$$\Delta p_t = -\underset{(0.017)}{0.010} + v_t$$

$$v_t + \underset{(0.169)}{0.982} v_{t-1} = \underset{(0.164)}{\varepsilon_t} + \underset{(\ldots)}{\varepsilon_{t-1}}$$

Rice: ARIMA(8,1,1), 88 DF , **Ljung-Box Q(10):** 0.801 (0.371), **Z**=-2.117

$$\Delta p_t = -\underset{(0.006)}{0.013} + v_t$$

$$\begin{aligned} v_t - \underset{(0.184)}{0.970} v_{t-1} + \underset{(0.147)}{0.575} v_{t-2} - \underset{(0.171)}{0.319} v_{t-3} + \underset{(0.162)}{0.290} v_{t-4} \\ - \underset{(0.165)}{0.223} v_{t-5} + \underset{(0.157)}{0.187} v_{t-6} - \underset{(0.144)}{0.233} v_{t-7} + \underset{(0.107)}{0.269} v_{t-8} \\ = \underset{(0.164)}{\varepsilon_t} - \underset{(0.172)}{0.791} \varepsilon_{t-1} \end{aligned}$$

[1] On this basis autocorrelated residuals are a possible problem in the estimated equations for Cotton, Jute and possibly Wheat (with 7.7%).

Wheat: ARIMA(5,1,1), 91 DF , **Ljung-Box Q(8):** 5.136 (0.077), **Z**=2.986

$$\Delta p_t = -\underset{(0.005)}{0.009} + v_t$$

$$v_t - \underset{(0.146)}{1.063} v_{t-1} + \underset{(0.143)}{0.364} v_{t-2} - \underset{(0.150)}{0.272} v_{t-3} + \underset{(0.144)}{0.395} v_{t-4} - \underset{(0.123)}{0.300} v_{t-5}$$

$$= \underset{(0.156)}{\varepsilon_t} - \underset{(n.a.)}{\varepsilon_{t-1}}$$

Maize: ARIMA(2,1,1), 94 DF , **Ljung-Box Q(8):** 3.295 (0.655), **Z**=-1.793

$$\Delta p_t = -\underset{(0.007)}{0.010} + v_t$$

$$v_t - \underset{(0.142)}{0.567} v_{t-1} + \underset{(0.112)}{0.231} v_{t-2} = \underset{(0.206)}{\varepsilon_t} - \underset{(0.116)}{0.779} \varepsilon_{t-1}$$

Sugar: ARIMA(5,1,1), 91 DF , **Ljung-Box Q(8):** 2.089 (0.352)1.8735 **Z**=1.736

$$\Delta p_t = -\underset{(0.018)}{0.012} + v_t$$

$$v_t + \underset{(0.308)}{0.649} v_{t-1} + \underset{(0.119)}{0.432} v_{t-2} + \underset{(0.158)}{0.329} v_{t-3} + \underset{(0.118)}{0.235} v_{t-4} + \underset{(0.101)}{0.295} v_{t-5}$$

$$= \underset{(0.321)}{\varepsilon_t} + \underset{(0.318)}{0.592} \varepsilon_{t-1}$$

Beef: ARIMA(0,1,1), 96 DF , **Ljung-Box Q(12):** 7.272 (0.401), **Z**= *n.a.*

$$\Delta p_t = \underset{(0.022)}{0.008} + v_t$$

$$v_t = \underset{(0.207)}{\varepsilon_t} + \underset{(0.102)}{0.055} \varepsilon_{t-1}$$

Lamb: ARIMA(5,1,1), 91 DF , **Ljung-Box Q(8):** 1.895 (0.388), **Z**=-3.883

$$\Delta p_t = \underset{(0.008)}{0.018} + v_t$$

$$v_t - \underset{(0.120)}{0.957} v_{t-1} + \underset{(0.130)}{0.030} v_{t-2} - \underset{(0.131)}{0.032} v_{t-3} - \underset{(0.132)}{0.322} v_{t-4} + \underset{(0.098)}{0.390} v_{t-5} = \underset{(0.197)}{\varepsilon_t} - \underset{(...)}{\varepsilon_{t-1}}$$

Bananas: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 3.894 (0.691), **Z**=2.300

$$\Delta p_t = \underset{(0.010)}{0.000} + v_t$$

$$v_t + \underset{(0.661)}{0.430} v_{t-1} = \underset{(0.092)}{\varepsilon_t} + \underset{(0.618)}{0.538} \varepsilon_{t-1}$$

Palm Oil: ARIMA(3,1,1), 93 DF , **Ljung-Box Q(8):** 4.648 (0.325), **Z**=3.306

$$\Delta p_t = -\underset{(0.016)}{0.007} + v_t$$

$$v_t - \underset{(0.316)}{1.032} v_{t-1} + \underset{(0.139)}{0.402} v_{t-2} - \underset{(0.169)}{0.332} v_{t-3} = \underset{(0.214)}{\varepsilon_t} - \underset{(...)}{\varepsilon_{t-1}}$$

Cotton: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 17.128 (0.009), **Z**=4.549

$$\Delta p_t = -\underset{(0.009)}{0.009} + v_t$$

$$v_t - \underset{(0.229)}{0.584} v_{t-1} = \underset{(0.164)}{\varepsilon_t} - \underset{(0.177)}{0.783} \varepsilon_{t-1}$$

Jute: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 12.149 (0.059), **Z**=5.014

$$\Delta p_t = -\underset{(0.011)}{0.009} + v_t$$

$$v_t - \underset{(0.205)}{0.616} v_{t-1} = \underset{(0.223)}{\varepsilon_t} - \underset{(0.153)}{0.818} \varepsilon_{t-1}$$

Wool: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 5.273 (0.509), **Z**=4.290

$$\Delta p_t = -\underset{(0.008)}{0.015} + v_t$$

$$v_t - \underset{(0.174)}{0.528} v_{t-1} = \underset{(0.190)}{\varepsilon_t} - \underset{(0.121)}{0.817} \varepsilon_{t-1}$$

Tobacco: ARIMA(5,1,1), 90 DF , **Ljung-Box Q(8):** 0.636(0.728), **Z**=1.874

$$\Delta p_t = \underset{(0.012)}{0.004} + v_t$$

$$v_t + \underset{(0.358)}{0.665} v_{t-1} - \underset{(0.126)}{0.056} v_{t-2} + \underset{(0.126)}{0.008} v_{t-3} + \underset{(0.127)}{0.242} v_{t-4} + \underset{(0.108)}{0.268} v_{t-5}$$

$$= \underset{(0.140)}{\varepsilon_t} + \underset{(0.364)}{0.706} \varepsilon_{t-1}$$

Rubber: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 7.060 (0.315), **Z**=9.165

$$\Delta p_t = -\underset{(0.009)}{0.030} + v_t$$

$$v_t - \underset{(0.137)}{0.921} v_{t-1} = \underset{(0.281)}{\varepsilon_t} - \underset{(\ldots)}{\varepsilon_{t-1}}$$

Timber: ARIMA(3,1,1), 93 DF , **Ljung-Box Q(8):** 5.856 (0.210), **Z**=2.023

$$\Delta p_t = \underset{(0.012)}{0.010} + v_t$$

$$v_t + \underset{(0.186)}{1.105} v_{t-1} + \underset{(0.153)}{0.339} v_{t-2} + \underset{(0.124)}{0.203} v_{t-3} = \underset{(160)}{\varepsilon_t} + \underset{(\ldots)}{\varepsilon_{t-1}}$$

Copper: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 6.001 (0.423), **Z**=2.007

$$\Delta p_t = -\underset{(0.021)}{0.009} + v_t$$

$$v_t + \underset{(0.744)}{0.400} v_{t-1} = \underset{(0.190)}{\varepsilon_t} + \underset{(0.704)}{0.504} \varepsilon_{t-1}$$

Aluminium: ARIMA(3,1,1), 93 DF , **Ljung-Box Q(8):** 5.422 (0.247), **Z**=2.867

$$\Delta p_t = -\underset{(0.009)}{0.019} + v_t$$

$$v_t - \underset{(0.246)}{1.187} v_{t-1} + \underset{(0.159)}{0.522} v_{t-2} - \underset{(0.139)}{0.288} v_{t-3} = \underset{(0.158)}{\varepsilon_t} - \underset{(\ldots)}{\varepsilon_{t-1}}$$

Tin: ARIMA(0,1,1), 96 DF , **Ljung-Box Q(8):** 4.126 (0.765), **Z**= *n.a.*

$$\Delta p_t = -\underset{(0.021)}{0.004} + v_t$$

$$v_t = \underset{(0.191)}{\varepsilon_t} + \underset{(0.102)}{0.066} \varepsilon_{t-1}$$

Silver: ARIMA(3,1,1), 93 DF , **Ljung-Box Q(8):** 5.173 (0.270), **Z**=2.462

$$\Delta p_t = -\underset{(0.015)}{0.003} + v_t$$

$$v_t + \underset{(0.135)}{0.896} v_{t-1} + \underset{(0.134)}{0.269} v_{t-2} + \underset{(0.114)}{0.256} v_{t-3} = \underset{(0.181)}{\varepsilon_t} + \underset{(0.091)}{0.968} \varepsilon_{t-1}$$

Lead: ARIMA(1,1,1), 95 DF , **Ljung-Box Q(8):** 4.972 (0.547), Z=8.467

$$\Delta p_t = -\underset{(0.005)}{0.005} + v_t$$

$$v_t - \underset{(0.107)}{0.851} v_{t-1} = \underset{(0.181)}{\varepsilon_t} - \underset{(\ldots)}{\varepsilon_{t-1}}$$

Zinc: ARIMA(2,1,1), 94 DF , **Ljung-Box Q(8):** 3.890 (0.565), Z=-2.026

$$\Delta p_t = \underset{(0.002)}{0.001} + v_t$$

$$v_t - \underset{(0.102)}{0.797} v_{t-1} + \underset{(0.101)}{0.204} v_{t-2} = \underset{(0.198)}{\varepsilon_t} - \underset{(\ldots)}{\varepsilon_{t-1}}$$

Appendix IV.iii. ADF Test Results Obtained Using Maximum Likelihood Estimation

For the modified pre-test method described in Sun and Pantula (1999) ADF tests with constant and trend term were calculated using Maximum Likelihood estimation, again employing the exact Maximum Likelihood routine for GAUSS which has been used in the remainder of this study. As for the OLS case, ADF testing equations were calculated with an initial number of five lags. Lagged terms were then eliminated where the last lagged term was insignificant at the asymptotic critical value of ±1.96. The final ADF testing equations obtained from this process are reported below (standard errors are given in parentheses below the estimated coefficient values) the ADF test statistic τ is listed separately.

Beef:

$$\Delta p_t = -\underset{(0.089)}{0.119} + \underset{(0.001)}{0.001}t - \underset{(0.045)}{0.085}p_{t-1} + \underset{(0.204)}{\varepsilon_t}$$

$$\tau = -1.876$$

Cotton:

$$\Delta p_t = \underset{(0.059)}{0.115} - \underset{(0.001)}{0.002}t - \underset{(0.070)}{0.128}p_{t-1} - \underset{(0.107)}{0.145}\Delta p_{t-1} - \underset{(0.106)}{0.455}\Delta p_{t-2}$$
$$-\underset{(0.112)}{0.287}\Delta p_{t-3} - \underset{(0.116)}{0.256}\Delta p_{t-4} + \underset{(0.155)}{\varepsilon_t}$$

$$\tau = 1.835$$

Lead:

$$\Delta p_t = \underset{(0.039)}{0.006} - \underset{(0.001)}{0.002}t - \underset{(0.083)}{0.175}p_{t-1} - \underset{(0.107)}{0.130}\Delta p_{t-1}$$
$$-\underset{(0.108)}{0.188}\Delta p_{t-2} - \underset{(0.110)}{0.303}\Delta p_{t-3} - \underset{(0.115)}{0.296}\Delta p_{t-4} + \underset{(0.184)}{\varepsilon_t}$$

$\tau = -2.105$

Maize:

$$\Delta p_t = \underset{(0.095)}{0.216} - \underset{(0.001)}{0.003}t - \underset{(0.103)}{0.227}p_{t-1} - \underset{(0.108)}{0.256}\Delta p_{t-1} - \underset{(0.110)}{0.465}\Delta p_{t-2}$$
$$-\underset{(0.119)}{0.316}\Delta p_{t-3} - \underset{(0.120)}{0.466}\Delta p_{t-4} - \underset{(0.130)}{0.310}\Delta p_{t-5} + \underset{(0.201)}{\varepsilon_t}$$

$\tau = -2.210$

Rubber:

$$\Delta p_t = \underset{(0.153)}{0.393} - \underset{(0.002)}{0.006}t - \underset{(0.074)}{0.228}p_{t-1} - \underset{(0.106)}{0.072}\Delta p_{t-1} - \underset{(0.106)}{0.236}\Delta p_{t-2} - \underset{(0.108)}{0.220}\Delta p_{t-3} + \underset{(0.279)}{\varepsilon_t}$$

$\tau = -3.104$

Timber:

$$\Delta p_t = -\underset{(0.123)}{0.384} + \underset{(0.002)}{0.005}t - \underset{(0.120)}{0.427}p_{t-1} - \underset{(0.111)}{0.255}\Delta p_{t-1} - \underset{(0.114)}{0.403}\Delta p_{t-2}$$
$$-\underset{(0.121)}{0.196}\Delta p_{t-3} - \underset{(0.124)}{0.363}\Delta p_{t-4} - \underset{(0.132)}{0.428}\Delta p_{t-5} + \underset{(0.154)}{\varepsilon_t}$$

$\tau = -3.550$

Wool:

$$\Delta p_t = \underset{(0.102)}{0.236} - \underset{(0.001)}{0.004}t - \underset{(0.073)}{0.155}p_t - \underset{(0.104)}{0.233}\Delta p_{t-1} - \underset{(0.107)}{0.469}\Delta p_{t-2}$$
$$-\underset{(0.113)}{0.237}\Delta p_{t-3} - \underset{(0.114)}{0.407}\Delta p_{t-4} + \underset{(0.184)}{\varepsilon_t}$$

$\tau = -2.137$

Appendix IV.iv. Estimation Results for the Price Series for Hides.

The series of relative prices for Hides has been omitted from the major part of this study, since the series has not been updated after 1995. Given the performance of this price series in the Vogelsang Test, however, ARIMA models selected by SBC have been estimated for comparative purposes. The Results are given here (with standard errors in parentheses). As in previous appendeces, the Ljung Box Q statistic is reported for 12 autocorrelations with P-values given in parentheses.

For the I(0) model:

Model: ARIMA (1,0,0)
SBC: 15.983 , **Ljung-Box Q(12):** 17.022 (0.107)
Degrees of Freedom: 93

$$p_t = \underset{(0.133)}{0.635} - \underset{(0.002)}{0.012}t + u_t$$

$$u_t - \underset{(0.080)}{0.633}\, u_{t-1} = \underset{(0.248)}{\varepsilon_t}$$

For the model in first differences:

Model: ARIMA (1,1,1)
SBC: 20.423 , **Ljung-Box Q(12):**18.390 (0.049)
Degrees of Freedom: 92

$$\Delta p_t = -\underset{(0.004)}{0.011} + v_t$$

$$v_t - \underset{(0.121)}{0.723}\, v_{t-1} = \underset{(0.251)}{\varepsilon_t} - \underset{(n.a.)}{\varepsilon_{t-1}}$$

Model in first differences selected by AIC:

Model: ARIMA(1,1,4)
AIC: 12.077, **Ljung-Box Q(12):** 9.040 (0.250)
Degrees of Freedom: 89

$$\Delta p_t = -\underset{(0.004)}{0.011} + v_t$$

$$v_t - \underset{(0.252)}{0.617}\, v_{t-1} = \underset{(0.246)}{\varepsilon_t} - \underset{(n.a.)}{0.902}\, \varepsilon_{t-1} - \underset{(n.a.)}{0.126}\, \varepsilon_{t-2} + \underset{(n.a.)}{0.288}\, \varepsilon_{t-3} - \underset{(n.a.)}{0.261}\, \varepsilon_{t-4}$$

ADF test result (testing equation with standard errors in parentheses below the estimate, the value of the ADF test statistic τ is given separately):

$$\Delta p_t = \underset{(0.091)}{0.332} - \underset{(0.002)}{0.006}\, t - \underset{(0.109)}{0.463}\, p_{t-1} + \underset{(0.123)}{0.117}\, \Delta p_{t-1} + \underset{(0.115)}{0.003}\, \Delta p_{t-2} + \underset{(0.107)}{0.248}\, \Delta p_{t-3} + \underset{(0.245)}{\varepsilon_t}$$

$$\tau = -4.245$$

Chapter 5

Forecasts of Average Annual Commodity Prices Relative to MUV

Chapter 5: Forecasts of Average Annual Commodity Prices Relative to MUV

Against the background of the evidence presented in the preceding two chapters, it should be obvious that conclusions on the order of integration of a series as well as inference about the presence of trend or drift components are crucial for the selection of forecast models. Where there is substantial uncertainty about the correct inference regarding either the order of integration or the presence and magnitude of a trend it also appears necessary to evaluate the likely cost of relying on the wrong model.

This chapter will commence with a presentation of the relevant forecast alternatives in general terms. After assessing the forecast performance of different model alternatives in terms of their forecast errors as well as the likely cost of using the wrong model a final selection of the forecast model to be used for each commodity price series will be made.

5.1 Forecast alternatives for the commodity series in the sample

The forecast model alternatives under consideration are the stationary and difference stationary models selected by minimum SBC, including and excluding a trend or drift term. In those cases where inference on the drift term changes for more extensive parameterisations of the difference stationary alternative, I(1) models selected by AIC are considered in addition. Also, ARIMA(1,1,1) models are considered as an alternative to random walks.

5.1.1 The forecast alternatives

When forecasting price series for the models in the sample, the possibility of using forecast models without trend or drift was considered for all price series. In this case, the ARIMA(p,d,q) models identified by SBC previously were re-estimated without the trend or constant included. Forecasts can then be obtained by extrapolating from the model for future time periods. It is worth noting that in the case of a difference stationary model without drift the only information that can be incorporated into the forecasts, aside from the last value of the original data series, is based on lagged residuals extrapolated via the autoregressive and moving average terms. This immediately implies, in the case of a pure moving average model, that no information from the historical residual process can be used in forecasts more than q periods ahead, since the expected value of the white noise residual term is zero if the model is specified correctly. In the case of a random walk this is reflected by the fact that the last available observation of the original data series provides the best available forecast. Where forecasts are based on a model in levels without trend, the forecasts revert to the unconditional mean of the series eventually. In the case of an ARIMA(0,0,q) model this occurs after q periods, while in those cases where autoregressive components are incorporated into the model specification mean reversion takes place more gradually.

To obtain forecasts from difference stationary models with drift, the minimum SBC models with drift, as presented in chapter 3 were extrapolated to future periods. An exception to this way of proceeding was given in those cases where relative commodity price series were identified as containing a significant drift only after model selection was repeated using the Akaike Information Criterion (AIC). Models with drift or trend were only considered among the alternatives for those commodities where there is some indication that a drift might be present. (In practice this was the case for the following six commodities: Rice, Sugar, Rubber, Timber, Aluminium and Wheat.)

The corresponding forecasts in levels are obtained by forecasting from trendless models, *i.e.* ARMA models including a constant, p autoregressive and q moving average terms, where once more the models previously selected by SBC were re-estimated without trend. As has been pointed out above, this is tantamount to using the unconditional mean as a forecast in the long run. Finally, forecasts including a trend were considered for all those commodity price series where the trend coefficient estimate was at least statistically significant at the standard critical values in the I(0) ARIMA model. Since standard ARIMA forecasting procedures in general, and the 'forecast.src' routine in GAUSS among them, do not normally incorporate linear trends it has been considered most appropriate to compute the projections with trend in a two step procedure: 1. the linear trend vector was subtracted from the original data series to obtain a stationary data series as a basis for the Box-Jenkins forecast, *i.e.* the forecast was based on $\boldsymbol{p}^* = \boldsymbol{p} - \widehat{\beta}\boldsymbol{t}$, where $\boldsymbol{p}$ is a vector containing the original data series, $\boldsymbol{t}$ is the linear trend vector and the scalar $\widehat{\beta}$ is the estimated trend coefficient. 2. Box-Jenkins forecasts were then computed in the usual way (*cf.* Granger and Newbold (1986)), using the coefficient estimates from the minimum SBC ARIMA model in levels. The product of the estimated trend coefficient and the linear trend vector was then added back onto the forecast[1] and the upper and lower interval values for the forecast obtained.

Given the number of forecast alternatives still under consideration for each model, the use of some kind of selection criterion for the most appropriate forecast model is clearly needed. It has been attempted, as far as possible, to decide on the presence of a trend or drift component independently of *a priori* assumptions on the order of integration of the data generating process. With respect to forecasting, the impact of the assumed order of integration remains an issue though, in its own right as well as with respect to any remaining interdependence regarding inference on the trend or drift coefficient. In the first instance, subsection 5.12 will explore the results of a study on the impact of pre-testing on forecast performance when the presence and magnitude of the trend coefficient are known *a priori*.

5.1.2. Pre-testing and Forecast Performance

In Chapter three some of the problems surrounding unit root tests were addressed and because of these issues attempts at measuring the magnitude of trend or drift coefficients as well as confirming their statistical significance relied only partially -if at all- on *a priori* testing. Nevertheless, and in spite of all the above problems, it is reported by Diebold and Killian (2000) that basing forecast model selection on ADF pre-tests improves prediction results[2].

---

[1] The confidence interval for the forecasts obtained therefore does not take account of the standard error of the point estimate of the trend coefficient either.

[2] One difference with the analysis undertaken here in previous chapters is that Diebold and Killian focus explicitly on the implications for forecast performance when the existence of a trend term is known, *i.e.* the objective here is to identify the best forecast model rather than the correct identification of the data generating process *per se*.

Diebold and Kilian (*op. cit.*) compare the forecast performance of a consistently applied differenced model with that of a model selected by pre-testing for unit roots. The differenced model is ARIMA (0,1,0) with drift while alternatively unit root pre-testing is used to chose between either an ARIMA (1,0,0) model with trend or an ARIMA (0,1,0) model with drift. Diebold and Kilian used Monte Carlo simulations with 20,000 replications for the data generating process:

[5.1.1.] $$p_t = \alpha + \beta t + u_t,\ u_t - \phi u_{t-1} = \varepsilon_t,\ \varepsilon_t \sim N(0,\sigma^2)$$

which clearly corresponds to the model expressed in [4.1.10]. The parameter values used by Diebold and Kilian are $\alpha = 7.3707$, $\beta = 0.0065$, $\sigma = 0.01$, and $\phi = 0.5, 0.9, 0.97, 0.99$ or 1. The ratio of mean squared errors for both forecast alternatives was then computed and plotted for different sample sizes and forecast horizons. The authors find that for anything but autoregressive parameters very close to unity, model selection on the basis of pre-testing yields superior forecast results to the alternative of using differenced models as a default. For an autoregressive parameter value of 0.9 these differences are negligible for small sample sizes and at very short forecast horizons. For low autoregressive coefficient values ($\phi = 0.5$) the pattern of performance differences is reversed: for small samples the underperformance of the differenced model alternative relative to the forecast model selected by pre-testing actually increases. It is still true though that the performance difference is smaller at shorter forecast horizons.

It was attempted here to replicate the results of Diebold and Kilian for a time series sample with 100 observations in a Monte Carlo experiment with 20,000 replications[3]. The alternative autoregressive coefficient values used in the data generating process for the simulations were $\phi = 0.7, 0.8, 0,9$ and 1, with the model for the DGP as above in [5.1.1]. The coefficient values for the constant and the trend coefficient (a and β) are also those used by Diebold and Kilian.

Figure 5.1.1. below shows the ratio of mean squared errors from forecasts from uniformly applied differenced models to those from forecasts obtained from ARIMA models selected by pre-testing. The model alternatives were ARIMA (1,0,0) vs ARIMA (0,1,0), including a trend or drift term respectively. The unit root testing procedure was the augmented Dickey-Fuller test including trend, constant and up to five lagged differenced terms. As in chapter 3, the critical value for eliminating lagged differenced terms of the dependent variable was taken to be +/- 1.96. In each replication of the simulation the forecast error was computed as:

[5.1.2] $$e = p - p^F,$$

were p is the vector of original data, p^F the vector of forecasts over h=100 periods. The Mean squared errors were then calculated as the arithmetic mean of the sum of squares of the forecast errors from individual replications, *i.e.*:

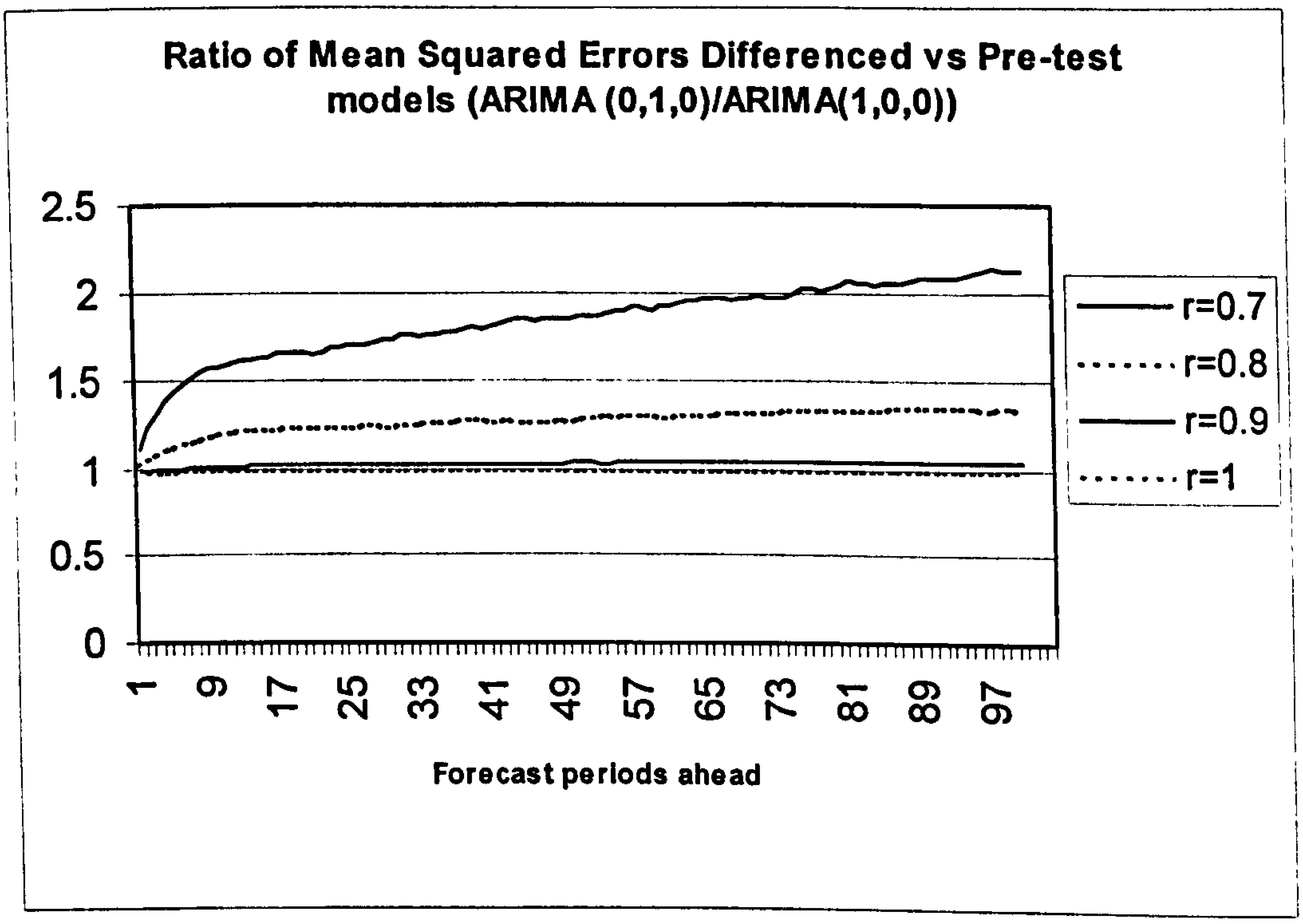
[5.1.3] $$\bar{e} = \begin{bmatrix} \frac{\sum_{i=1}^{n} e_{1,i}^2}{n} \\ \frac{\sum_{i=1}^{n} e_{2,i}^2}{n} \\ \vdots \\ \frac{\sum_{i=1}^{n} e_{h,i}^2}{n} \end{bmatrix}$$

---

[3] For a general description of the simulation methodology adopted here see Chapter 4. Simulations for forecasts were undertaken allowing 100 observations prior to the simulated data series for the construction of ARMA residual processes as in Chapter 4. The sample size for forecast model estimation was set at T=100 throughout. In each case the data series covered a further h=100 observations as a reference point for calculating forecast errors.

In [5.13] the subscript i refers to the current of the total of n replications, while $1,2...h$ denotes the forecast horizon for which the error is being computed.

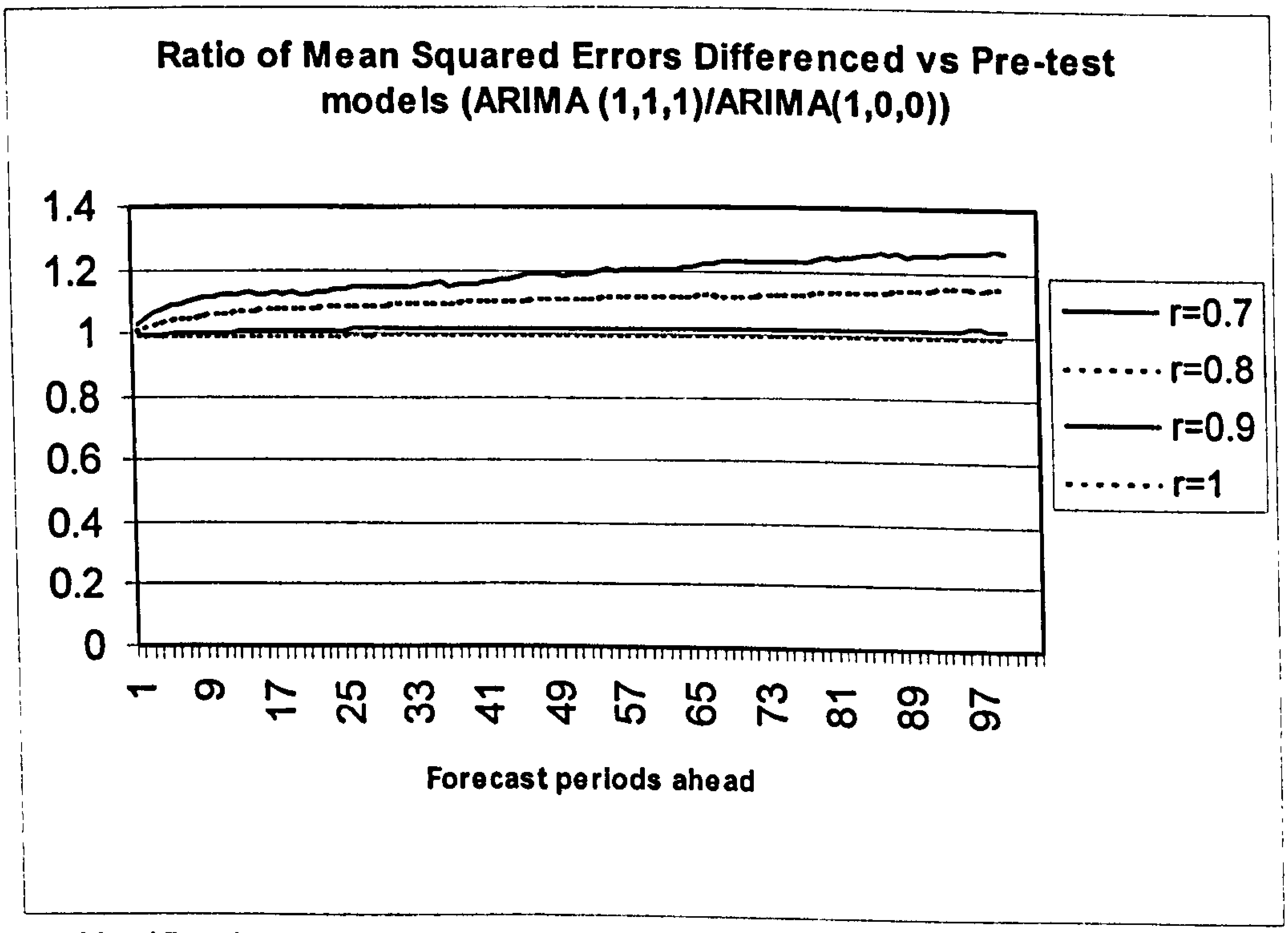
Figure 5.1.1 Simulation Evidence:



'r' identifies the autoregressive coefficient. The order of AR coefficient values in the legend coincides with the position of the respective line in the graph., *i.e.* the uppermost line is obtained for r=0.7 *etc.* Simulations were run over 20000 replications.

It can be seen that the Mean Squared Errors obtained from the forecasts based on ARIMA (0,1,0) models are consistently above those for the forecasts from models chosen by pre-testing. These differences are small for an autoregressive coefficient value of $\phi = 0.9$, but become larger for coefficient values of $\phi = 0.8$ and 0.7, as is to be expected given the low power of the unit root test in cases where the autoregressive coefficient is close to the stationarity boundary. In a similar comparison the model alternatives for the model selection by pre-testing were ARIMA(1,0,0) and ARIMA(1,1,1) while an ARIMA(1,1,1) model is consistently applied for forecasts from a differenced model. (The fitted models again contain a trend or drift term as appropriate.) This comparison is of interest in so far as it was demonstrated in chapter 4 that the successful detection of significant drift components becomes difficult in underparameterised difference stationary models. For the case of an ARIMA (1,0,0) data generating process this problem was shown to be attenuated by the use of ARIMA (1,1,1) models as the difference stationary alternative specification. For some data series an ARIMA (1,1,1) vs ARIMA (1,0,0) model specification will therefore be relevant alternatives. More generally, a similar problem arises where there is uncertainty between choosing either a model in levels or a difference stationary model identified by AIC.

Figure 5.1.2 Simulation Evidence:



'r' identifies the autoregressive coefficient. The order of AR coefficient values in the legend coincides with the position of the respective line in the graph, *i.e.* the uppermost line is obtained for r=0.7 *etc.* Simulations were run over 20000 replications

Figure 5.1.2. above shows that the general qualitative characteristics of the relative forecast performance of the two model selection techniques remain unaltered for the modified model alternatives. The extent to which the ARIMA (1,1,1) model underperforms the model selected through pre-testing is, however, smaller. The ratio of forecast errors is still higher for smaller values of the autoregressive coefficient, but the value of that ratio is now lower than the corresponding ratio in the case where the difference stationary alternative was a random walk with drift. What these simulation results give an indication of is that an ADF test can conceivably improve univariate forecast performance in spite of the problems of unit root tests outlined above. Some caution should be exercised in interpreting these results more generally, since no detailed information is available on how the results obtained would change for different model parameterisations. Another problem not considered here is the impact of structural instability in the data generating process on the quality of forecasts from models obtained by pre-testing and on the validity of the pre-test itself. A further limitation of the results presented by Diebold and Killian (*op. cit.*) is that simulations are undertaken to explore the performance of model alternatives in a case where the presence of a trend in the data generating process is known. Diebold also refers to Stock (1996) who confirms that forecast performance improvements result from the use of unit root tests in model selection when it is known that the data generating process does not contain a trend or drift component. It has been shown above, however, that the inference about the significance of trend or drift coefficient estimates is itself closely related to inferences about the order of integration when conventional

stationarity and significance testing procedures are being employed. Diebold and Killian's statement to the effect that pre-testing for unit roots improves the observed forecast performance on average, irrespective of whether the order of integration of the data generating process is correctly identified by the testing procedure should be seen against this background. The observed performance improvements should not be taken for granted if conclusions about the significance of the trend coefficient estimate have not been reached in a manner that does not directly depend on the assumed order of integration.

In the following, some additional simulation evidence is used to look into the impact of different assumptions regarding the order of integration and the presence of a trend or drift on the basis of different alternatives for the data generating process.

5.1.3. Differences in the forecast performance of alternative models

A study by Clements and Hendry (2001).

A more general study regarding the impact of the assumed order of integration on the magnitude of the prediction error, but abstracting from the issue of pre-testing, has been conducted by Clements and Hendry (2001). The authors investigate the role of *a priori* assumptions on the order of integration allowing for parameter uncertainty as well as for the case of a known DGP. They further consider the case of a fixed sample size as well as forecasts from samples of increasing size. A forecast horizon length of up to h=100 as in the present study has been considered. Clements and Hendry (*op. cit.*) furthermore base their results on theoretical calculations as well as Monte Carlo simulations. The authors confirm that, in general, mistaken assumptions on the order of integration of the data generating process lead to higher forecast errors regardless of whether the parameter values are known or have to be estimated and regardless of whether the sample size is fixed or increasing.

Clements and Hendry do not, however, look into the consequences of omitting or including a trend coefficient in the estimated model when the DGP does or does not contain a trend or drift component. Neither do they elaborate on the interdependence between inference on the order of integration and the presence of a trend or drift term. In addition to the impact of different *a priori* assumptions about the order of integration, the issue of inference regarding trend or drift coefficients is taken up in the simulation experiments on forecast performance described below.

Simulation results on forecast performance and model parameterisation

Comparing the performance difference between models with or without a trend or drift term one should expect the importance of including a trend or drift component to become more important over longer time horizons. Assumptions on the order of integration are relevant in so far as a correct representation of the data generating process should lead to better forecasts on average. The simulation results outlined below will give some indication of the importance of selecting between different model alternatives, for the trend coefficient values considered in the present case and over a forecast horizon length of up to 100 periods.

Among the results obtained, the differences in point forecasts from models in levels and difference stationary models can be substantial. The point estimates of the trend coefficients obtained are however, often of small magnitude, (the implications of this will be discussed in more detail below). It is not clear then how substantial the impact of including a trend component in the forecast model is under different circumstances and how big the impact on forecast accuracy would be.

Although the forecast results themselves give some indication of the differences in the point forecasts obtained from different model alternatives, it is not necessarily obvious how the respective models would perform in terms of forecast errors. To explore this issue further, simulations were conducted for ARIMA(1,0,0) models according to [5.1.1] with $a=1$ and where β took values between 0 and 0.1 increasing by values of 0.01. Simulations were conducted over 20,000 replications fitting ARIMA(1,0,0) models as well as ARIMA(0,1,1) models with and without trend or drift as appropriate and allowing for autoregressive coefficient values of $\phi=0.8, 0.9$ in the data generating process. As before, the sample size was fixed at T=100 with forecast horizons up to h=100. Mean squared forecast errors where then computed as in [5.1.2]-[5.1.3] above.

For comparative purposes, further simulations were then conducted for difference stationary processes of the form:

[5.1.4.] $$p_t = p_{t-1} + \beta + v_t \text{ , } v_t = \varepsilon_t - \theta\varepsilon_{t-1}$$

where the values used for θ were $\theta = 0.1, 0.2$. The values for the drift parameters were the same as above and simulations were run over 20,000 replications, again fitting ARIMA(1,0,0) and ARIMA(0,1,1) models with and without drift. Mean squared forecast errors were again calculated as above, in the simulations on pre-testing. The retained mean squared errors can be used in a number of ways to evaluate the forecast performance of different model alternatives. The results obtained should serve the purpose of characterising the performance of different forecast model alternatives under different scenarios of shock persistence.

Computing the mean squared forecast error as above in [5.1.3] it is possible to compare the forecast errors arising from different model alternatives fitted to the same generating model. For the trend stationary model according to [5.1.1], the following ratios of mean squared forecast errors were computed to evaluate the performance of various forecast alternatives relative to the generating model:

1. *Allowing for mistaken conclusions on the presence or absence of a trend term* by comparing the mean squared errors of models in levels with and without trend. Generally, this is the ratio of the mean squared forecast error of forecasts from the trendless model in levels relative to the forecast error of forecasts from a model in levels including trend (*i.e.* the model corresponding to the DGP). The one exception to this is the case where the trend coefficient in the generating process takes a value of zero. In this case, the mean squared errors of forecasts from a model with trend are indicative of the performance of the counterfactual model, while those from a stationary model excluding trend, are representative of forecasts from a model corresponding to the DGP.
2. *Considering a misspecification of the order of integration of the data generating process.* Allowing alternatively for the possibility of modelling the data series on the basis of a mistaken inference about the order of integration of the series, the second statistic is the ratio of the mean squared forecast error of

forecasts from difference stationary models including a drift component divided by the mean squared forecast error of forecasts from a stationary model including trend. Again, there is a difference in the case where the trend coefficient is set to equal zero in the data generating process and where correspondingly, the comparison is made between the forecast errors of a trendless difference stationary model and the forecast errors from a trendless model in levels.

3. *Considering the impact of mistaken inference regarding both the order of integration of the data generating process and the presence of a trend:* The third ratio is computed allowing for misspecifications not only of the order of integration but also for mistaken inference regarding the presence of a trend or drift term. Where $\beta = 0$ by construction, the comparison is made between forecast errors from an I(1) model with drift and a model in levels without trend. Otherwise, the ratio is computed from the mean squared forecast errors from driftless difference stationary models divided by the mean squared forecast error from a trend stationary model.

A similar comparison was made for the difference stationary data generating process outlined in [5.1.4]. Here the ratios used are as follows:

1. *Allowing for mistaken conclusions on the presence or absence of a drift term* by comparing the mean squared errors of models in first differences with and without drift. The mean squared forecast errors in the denominator are those corresponding to forecasts from a model with drift while the numerator is defined by the mean squared forecast errors from the counterfactual driftless

model. Again, the mean squared prediction errors in the numerator are those of the counterfactual model, where the data series are generated from a difference stationary model without drift.

2. *Considering a misspecification of the order of integration of the data generating process.* In this case the comparison is once more between models including trend or drift as above although the ratio is now inverted since the data generating process is difference stationary. For the case where $\beta = 0$, the ratio is taken between the mean squared forecast errors from trendless and driftless models. Accordingly the forecast error ratio is computed from the mean squared prediction errors from trend stationary models and difference stationary models with drift for cases with a non zero drift coefficient in the DGP.
3. *Considering the impact of mistaken inference regarding both the order of integration of the data generating process and the presence of a drift term:* As before, the third ratio is computed allowing for misspecifications not only of the order of integration but also for mistaken inference regarding the presence of a trend or drift term. Once more, the mean squared error ratio is inverted, where $\beta = 0$ by construction. In this case the comparison is made between forecast errors from an I(0) model with trend and a model in first differences without drift. Naturally, this third aspect is mainly relevant in the case of a zero drift coefficient in the data generating process since in this case the possibility of wrongly inferring the presence of a trend in a stationary forecast model has been shown, in Chapter 4, to be a likely outcome.

In the remaining cases the mean squared prediction errors in the numerator of the forecast error ratio arise from forecasts on the basis of a trendless stationary model while those in the denominator correspond to forecasts from an I(1) model with drift[4].

5.1.4. Simulation results for the forecast performance of various model alternatives

To obtain a general impression of how alternative model specifications perform under different scenarios, the three ratios defined above can be observed over increasing forecast horizons for the four alternative data generating processes considered here. Initially, the focus will be on the performance of different forecast models fitted to a stationary or trend stationary data generating process.

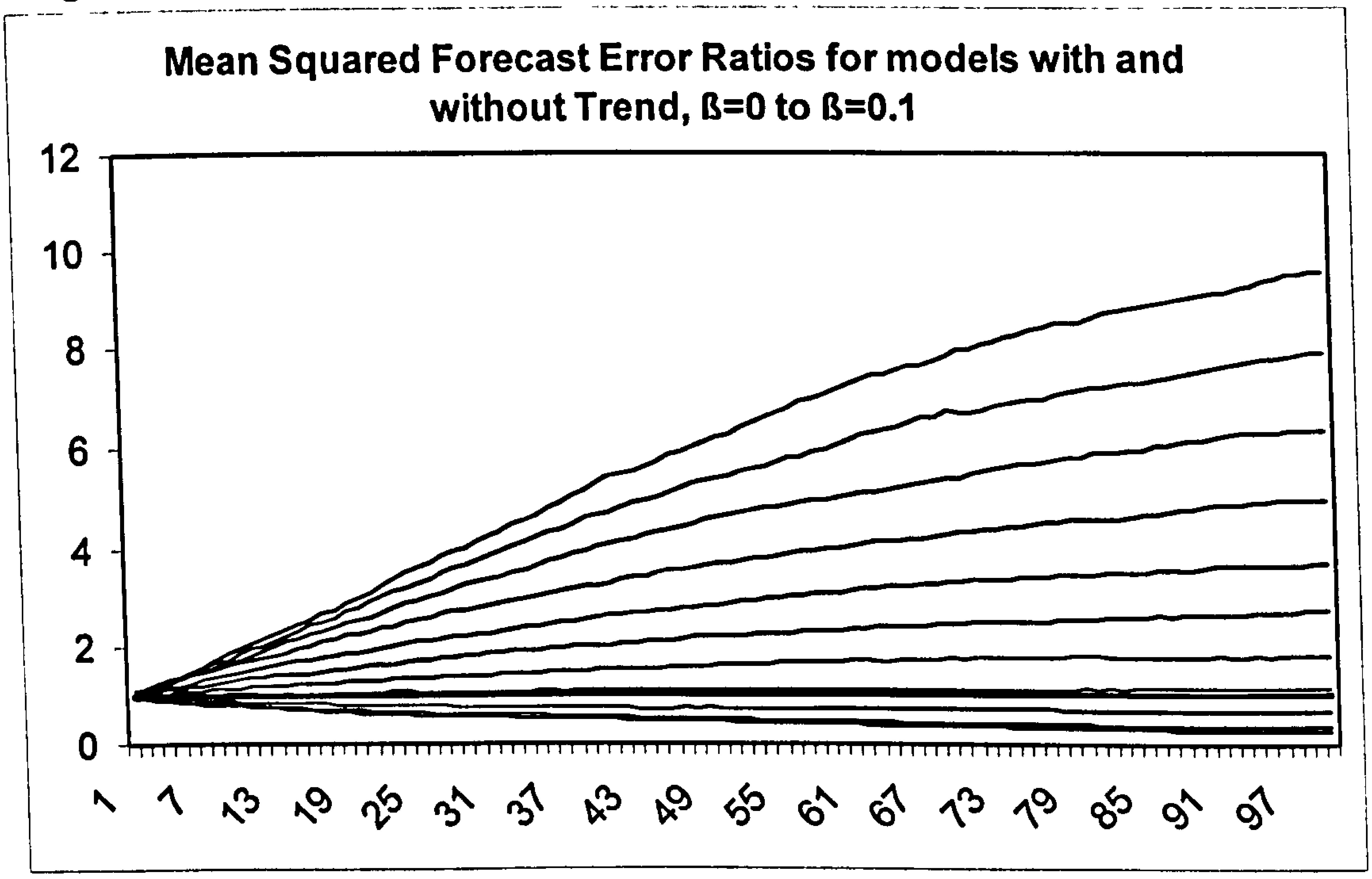
DGP = ARIMA(1,0,0) with $\phi = 0.9$

1. Allowing for erroneous inference on the presence of a trend term in the data generating process. Considering first the relative forecast performance of ARIMA(1,0,0) models with and without trend, when the trend coefficient in the data generating process is zero by construction, it can be seen from the ratio of mean squared forecast errors that models without trend consistently outperform

[4] Some of the simulation experiments mentioned in Clements and Hendry (2001) are similar to those reported here for the case of mistaken inference on the order of integration of the DGP only. However, there is one difference in the specification of the counterfactual model applied to various DGPs. When simulating a trend stationary DGP Clements and Hendry (*op. cit.*) rely on an AR(1) model with constant and trend, with an AR(1) coefficient value of 0.9 being closest to the experiment considered here. The counterfactual difference stationary model fitted is a random walk plus drift as opposed to the ARIMA(0,1,1) model used here. When simulating a difference stationary process, Clements and Hendry (*op. cit.*) use an ARIMA(0,1,1) model while fitting a counter factual trend stationary model with white noise error term, instead of allowing for autocorrelated residuals. It is not entirely clear why different model specifications have been used for generating and fitted counterfactual models for either order of integration.

models with trend over the whole of the forecast horizon. This general pattern persists for trend coefficient values of up to 0.02 in the DGP: in all these cases the ratio of the mean squared forecast error of models without trend to the mean squared forecast error for models with trend takes values below one over the entire forecast horizon. From trend coefficient values of 0.03 and above the opposite is the case: models including a trend term now yield superior forecast performance. This is illustrated in Figures 5.1.3 and 5.1.4 below where the ratios of mean squared forecast errors from forecast models with and without a trend term are plotted.

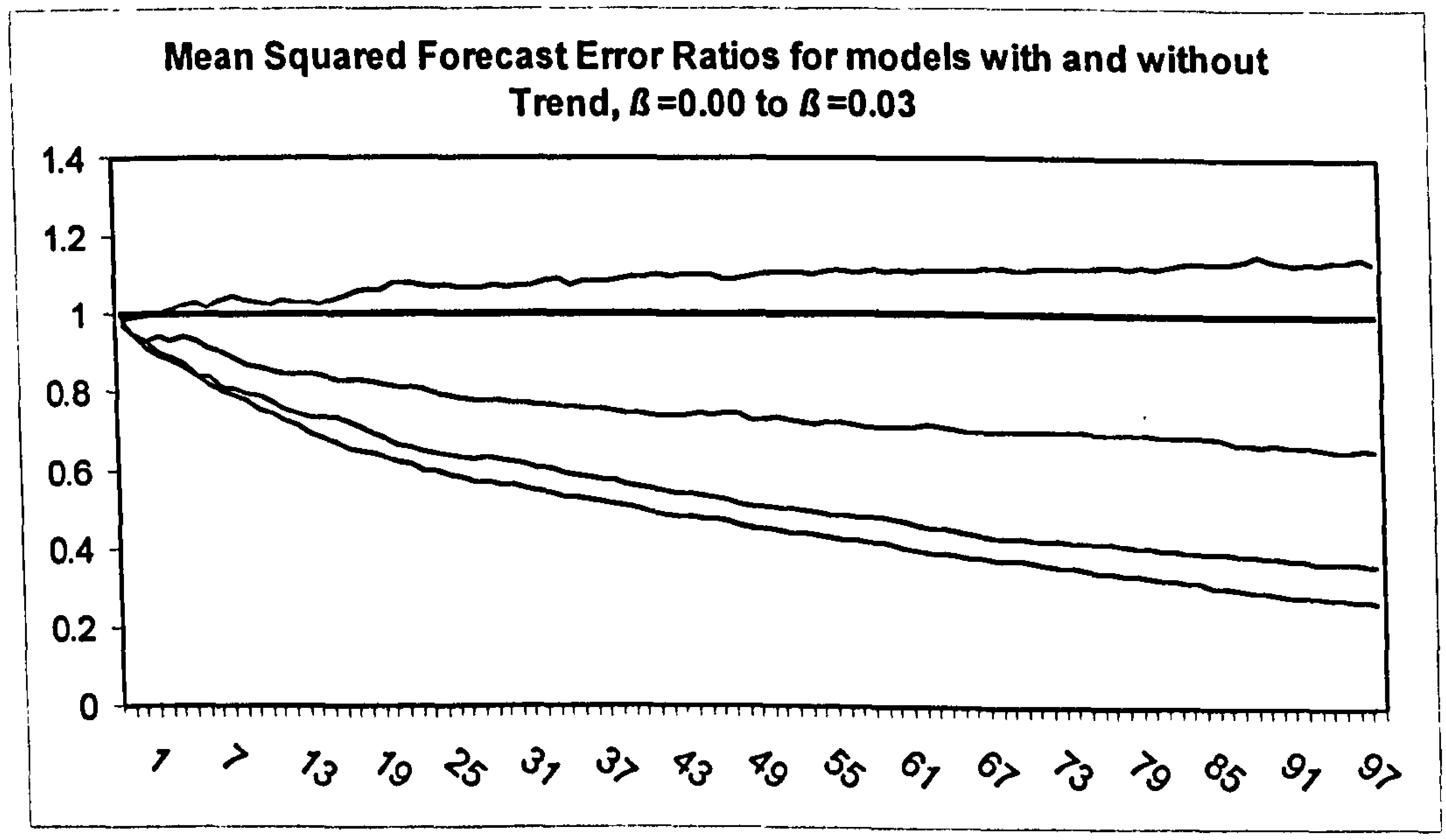
Figure 5.1.3 Simulation Evidence:



DGP is $p_t = a+\beta t+u_t,\ u_t-0.9u_{t-1}=\varepsilon_t$. Number of replications: 20000. Mean Squared Errors for models with trend in the denominator. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model. Higher positioned lines correspond to a higher value for the trend coefficient in the Data Generating Process.

The lowest line in Figure 5.1.3 corresponds to a data generating process with a zero trend coefficient. For all of the following lines a higher position of the line illustrating the ratio of mean squared errors corresponds to a higher trend coefficient in the data generating process, with trend coefficient values increasing by values of 0.01. Thus the line immediately following the lowest line illustrates the ratio of prediction mean squared errors for a trend coefficient of 0.01 in the DGP, the line above this corresponds to a trend coefficient value of 0.02 in the DGP *etc.* The next higher line, which is also the first one to be almost consistently positioned above the unity line corresponds to a generating trend coefficient value of 0.03. Graphs for the mean squared forecast error ratios obtained for trend coefficients of magnitude 0-0.03 in the DGP are illustrated separately in Figure 5.1.4.

Figure 5.1.4 Simulation Evidence:



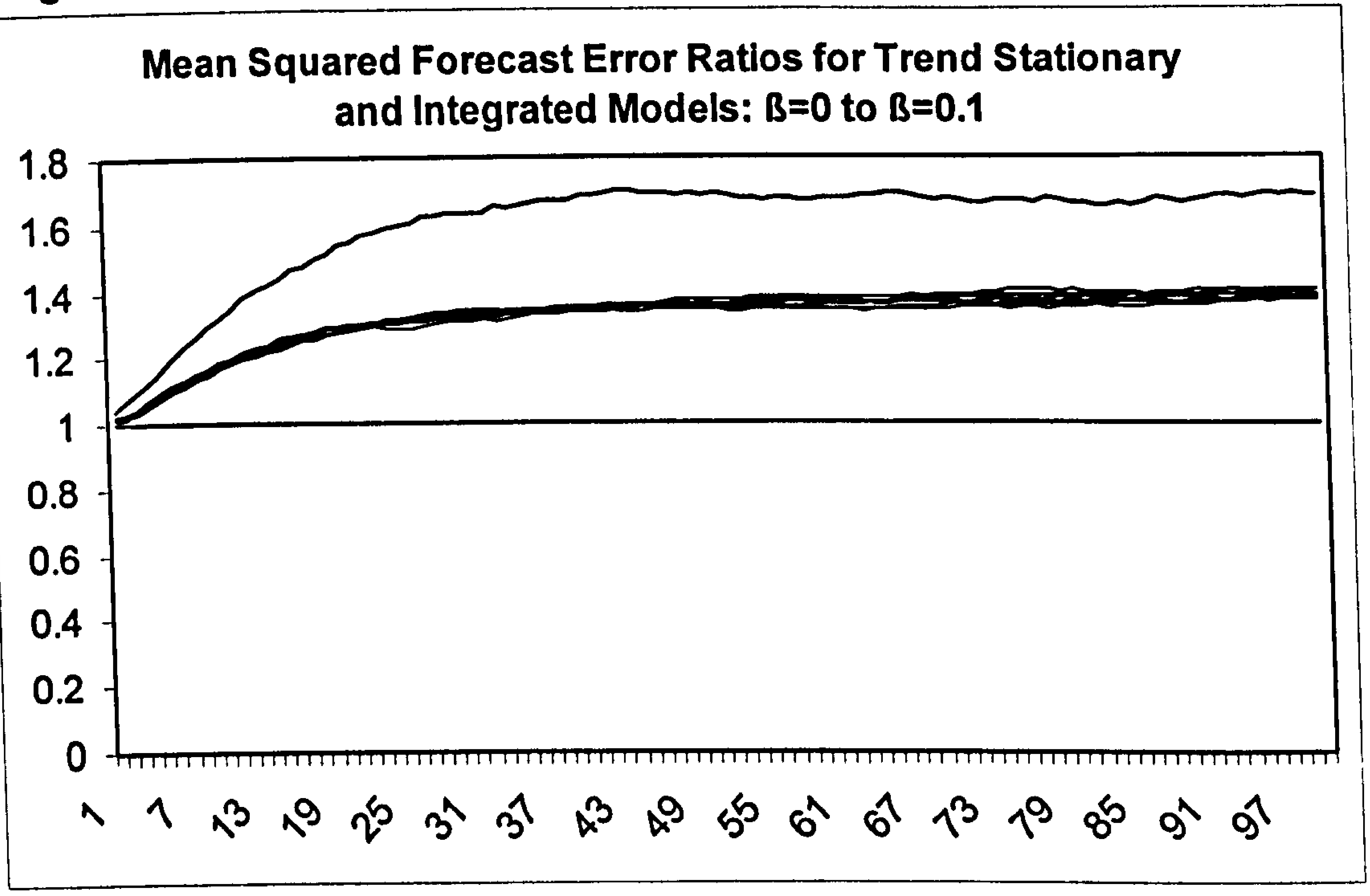
DGP is $p_t = a + \beta t + u_t$, $u_t - 0.9u_t = \varepsilon_t$. Number of replications: 20000. Mean Squared Errors for models with trend in the denominator. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model. Higher positioned lines correspond to a higher value for the trend coefficient in the Data Generating Process.

The lowest line shows the mean squared forecast error ratio for a zero trend coefficient. The line directly above corresponds to a generating trend coefficient value of $\beta = 0.01$, and the next higher to $\beta = 0.02$. It is readily seen that the mean squared forecast error of the counterfactual model increases relative to that of the correct forecast model, as the trend coefficient value rises. Up to a trend coefficient value of 0.02 in the DGP this is reflected by the line illustrating this ratio approaching the unity line. For a trend coefficient value of $\beta = 0.03$ the counterfactual model yields higher mean squared forecast errors for forecast horizons of h=4 and beyond.

2. ***Comparing forecasts from models where the order of integration has been inferred correctly with those where this was not the case.*** The ratio of mean squared forecast errors from models in first differences to the mean squared error of forecasts from the corresponding model in levels is here calculated for the case where the presence or absence of a trend or drift coefficient has been inferred correctly in accordance with the data generating process. Taking the ratio of mean squared forecast errors from forecasts without drift to the mean squared error of forecasts without trend, the forecast model in levels without trend is seen to have smaller forecast errors over the entire forecast horizon. The same general pattern is observed for models with trend or drift component. The comparison in this case is made taking the ratio of the mean squared forecast errors from an I(1) model with drift to those of an I(0) model with trend. Again, it is seen that lower forecast errors result if the order of integration of the data generating process is inferred correctly. This holds true for all the coefficient values considered and over the

whole 100 period forecast horizon. The superior performance of forecast models correctly identifying the order of integration of the DGP is shown in Figure 5.1.5. below, where the ratio of mean squared prediction errors from the counterfactual model over those corresponding to forecasts from the correct model specification is shown to consistently lie above one[5].

Figure 5.1.5 Simulation Evidence:



DGP is $p_t = \alpha + \beta t + u_t$, $u_t - 0.9u_t = \varepsilon_t$. Number of replications: 20000. Mean Squared Errors for models in levels are in the denominator. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model. For $\beta = 0$ the comparison is between forecasts from models without trend or drift.

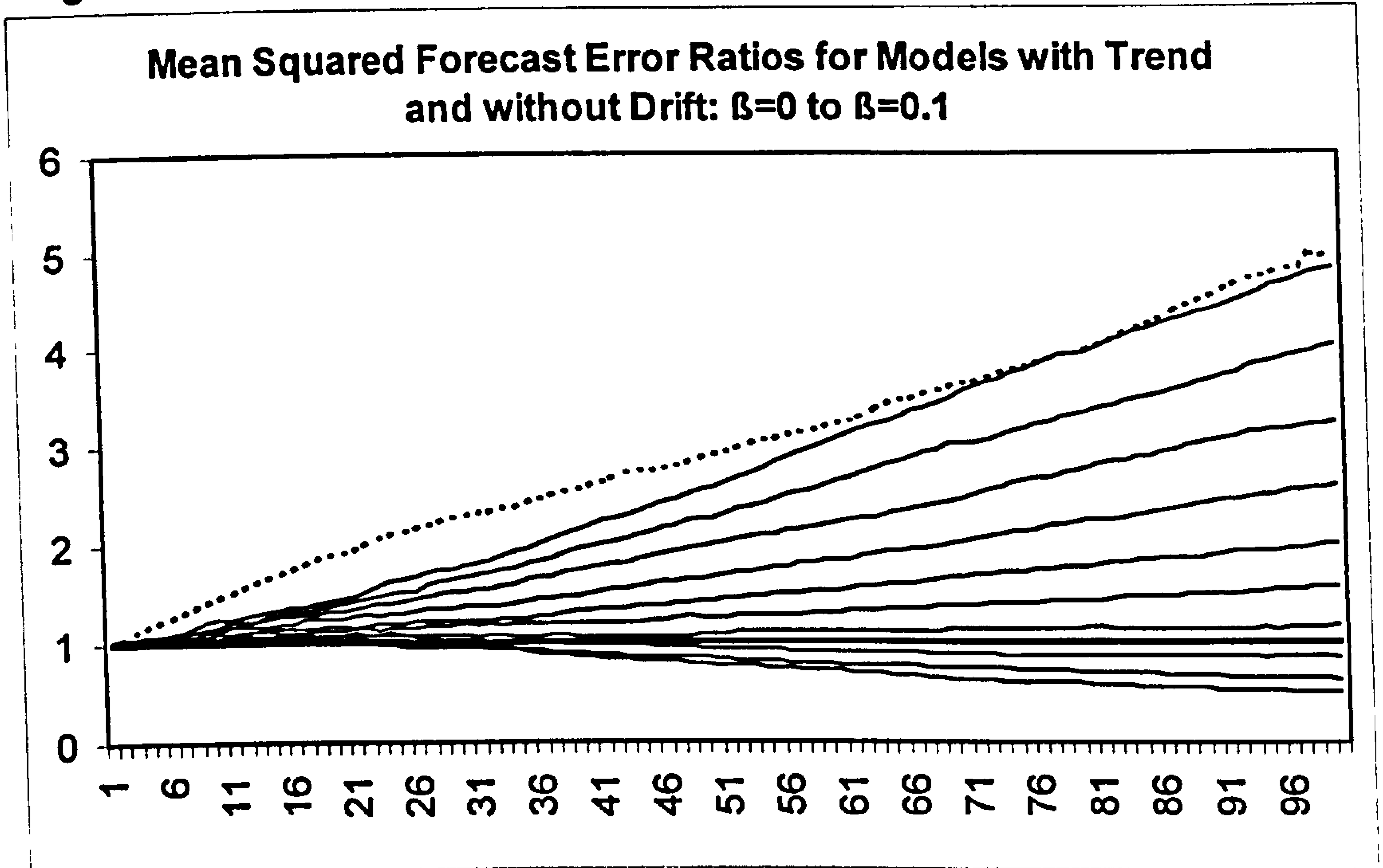
3. Considering mistaken inference on both the order of integration and the presence of a trend term. To allow for the possibility of basing the forecast model on mistaken conclusions about the order of integration as well as the presence of a trend component, forecasts were compared between an ARIMA(0,1,1) model

---

[5] Of course, the lines illustrating the mean squared forecast error ratios for different trend coefficient values in the DGP can not be distinguished individually, but the general pattern should be clear. The line in the highest position in Figure 5.1.5 corresponds to the case where $\beta = 0$.

without drift and the correct specification of an ARIMA(1,0,0) model with trend. (Clearly this is relevant as this is close to the forecast error scenario in those cases where the wrong order of integration is inferred for a data series and the drift coefficient is rejected as insignificant in consequence.) Again this comparison was made taking the ratio of the mean squared forecast error for the ARIMA(0,1,1) model excluding drift and the mean squared forecast error for the ARIMA(1,0,0) model with trend. The exception occurs again where the trend coefficient in the data generating process takes a value of zero so that the comparison is made between mean squared forecast errors from an I(1) model with drift and the mean squared forecast error from an ARIMA(1,0,0) model without trend. In this case, where $\beta = 0$ forecast errors are unambiguosly smaller when the correct model parameterisation is used as a basis for forecasts. For very small trend coefficient values, however, the situation is not as clear. The general pattern is illustrated below in Figure 5.1.6.

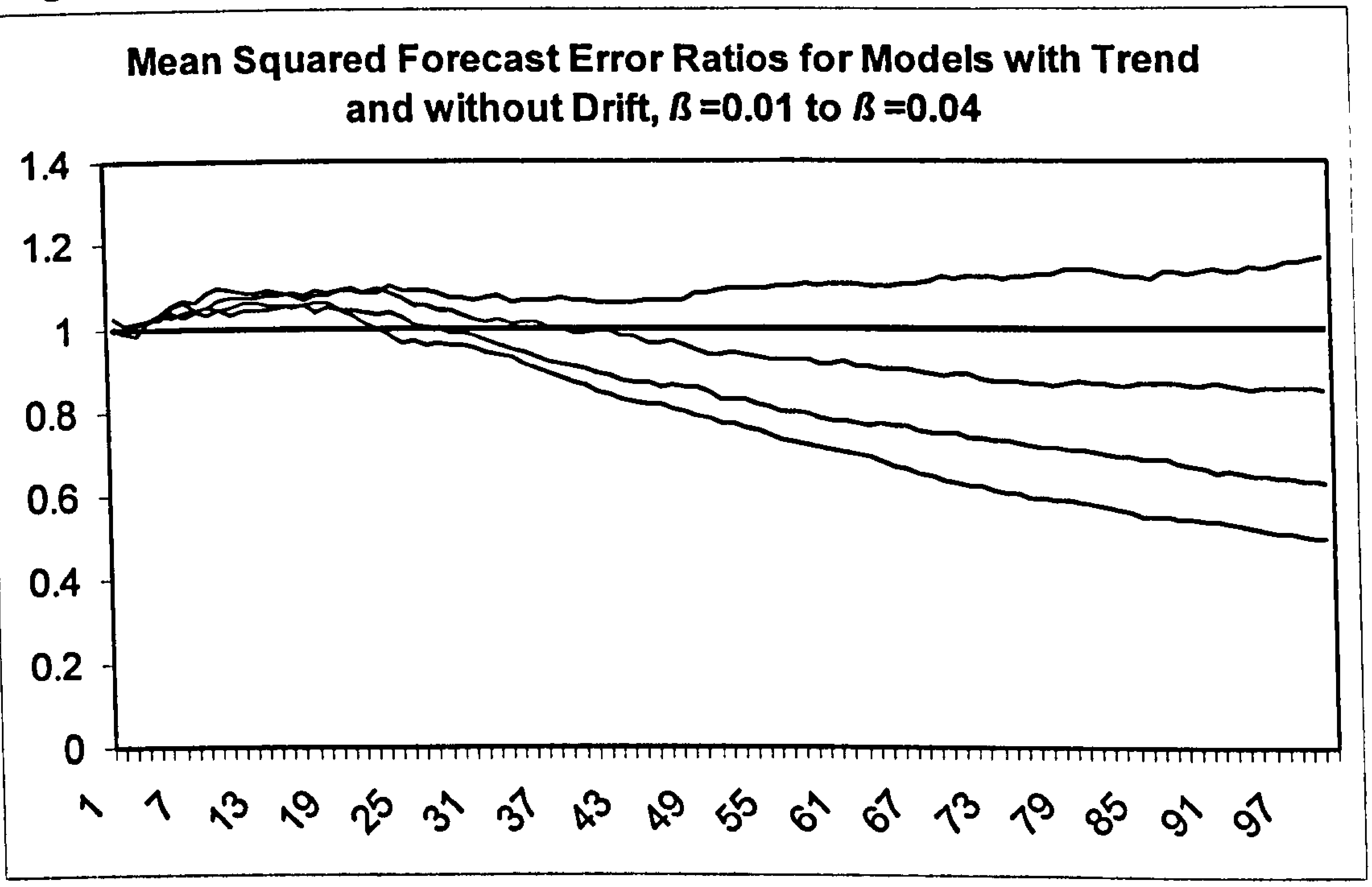
Figure 5.1.6 Simulation Evidence:



DGP is $p_t = a + \beta t + u_t$, $u_t - 0.9u_t = \varepsilon_t$. Number of replications: 20000. Mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

The dotted line in Figure 5.6 illustrates the mean squared prediction error ratio for the case where the trend coefficient in the DGP takes a zero value and the counterfactual model is difference stationary, allowing for a constant. The solid lines represent the cases of non zero trend coefficients in the DGP, increasing by values of 0.01 from $\beta = 0.01$, and with higher lines corresponding to higher values for the trend coefficient. It appears that for sufficiently large trend coefficient values, correctly including a trend tends to improve forecast performance consistently. For small trend coefficient values, however a better forecast performance can be inferred for the counterfactual model at long forecast horizons. This is illustrated by the lower lines showing mean squared forecast errors taking values below one at large values of h. This property of processes containing small trend components is shown separately in Figure 5.1.7.

Figure 5.1.7 Simulation Evidence:



DGP is $p_t = a + \beta t + u_t$, $u_t - 0.9u_t = \varepsilon_t$. Number of replications: 20000. Mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

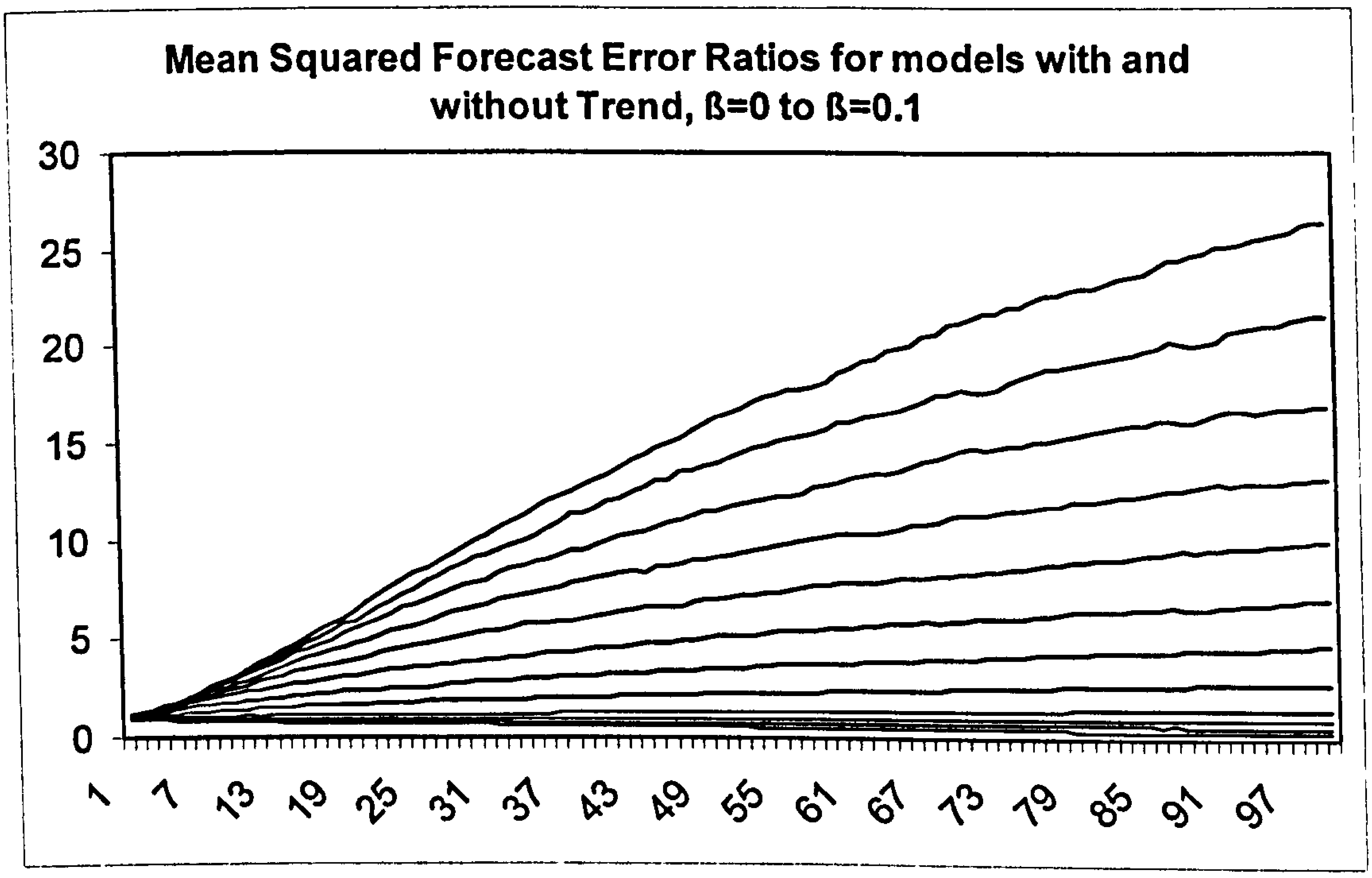
The lowest line corresponding to a trend coefficient value of $\beta = 0.01$, shows that in this case forecasts from a trendless difference stationary model have smaller forecast errors after a forecast horizon of 23 periods. This lengthens to h=28 forecast periods for $\beta = 0.02$, and h=37 for $\beta = 0.03$. (Here again, higher lines correspond to higher trend coefficient values.) As the trend coefficient reaches a magnitude of 0.04 and above, smaller forecast errors are associated with the forecast model parameterisation corresponding to the data generating process.

Thus, although the counterfactual model can yield superior forecasts at very low trend coefficient values and for long forecast horizons this does not come to bear for most of the duration of the 10 to 20 year horizons considered here.

DGP = ARIMA(1,0,0) with $\phi = 0.8$

1. Allowing for erroneous inferences on the presence of a trend term in the data generating process. The method of comparison used is as for the case of $\phi = 0.9$ and the results obtained are largely similar. One difference worth noting is, that forecast models which include a trend now yield superior forecast results for trend coefficients of $\beta \geq 0.02$ in the data generating process, while a value of $\beta = 0.03$ was needed in the case of $\phi = 0.9$. Figure 5.1.8 below shows, how the general pattern in this case is similar to the one observed for $\phi = 0.9$:

Figure 5.1.8 Simulation Evidence:

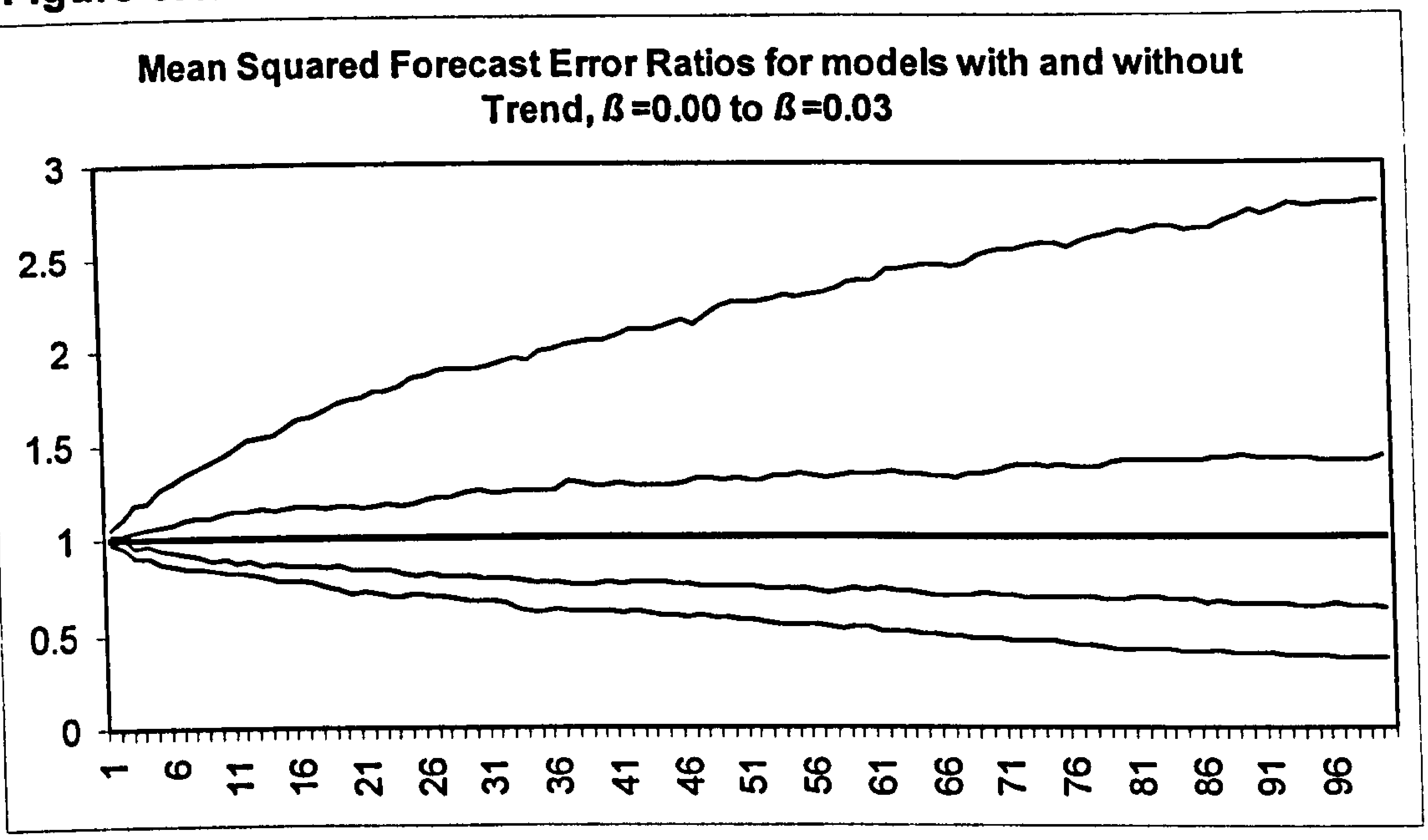


DGP is $p_t = a + \beta t + u_t$, $u_t - 0.8u_t = \varepsilon_t$. Number of replications: 20000. Except for the case where $\beta = 0$ mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

The case for low values of the trend coefficient in the DGP is illustrated separately in Figure 5.1.9. below.

Again, lower lines correspond to lower trend coefficient values in both figures 5.1.8 and 5.1.9. Models without trend now perform better than models with trend only for trend coefficient values of $\beta = 0$ and $\beta = 0.01$, at values of $\beta \geq 0.02$ models including a trend term have lower mean squared forecast errors than the counterfactual model.

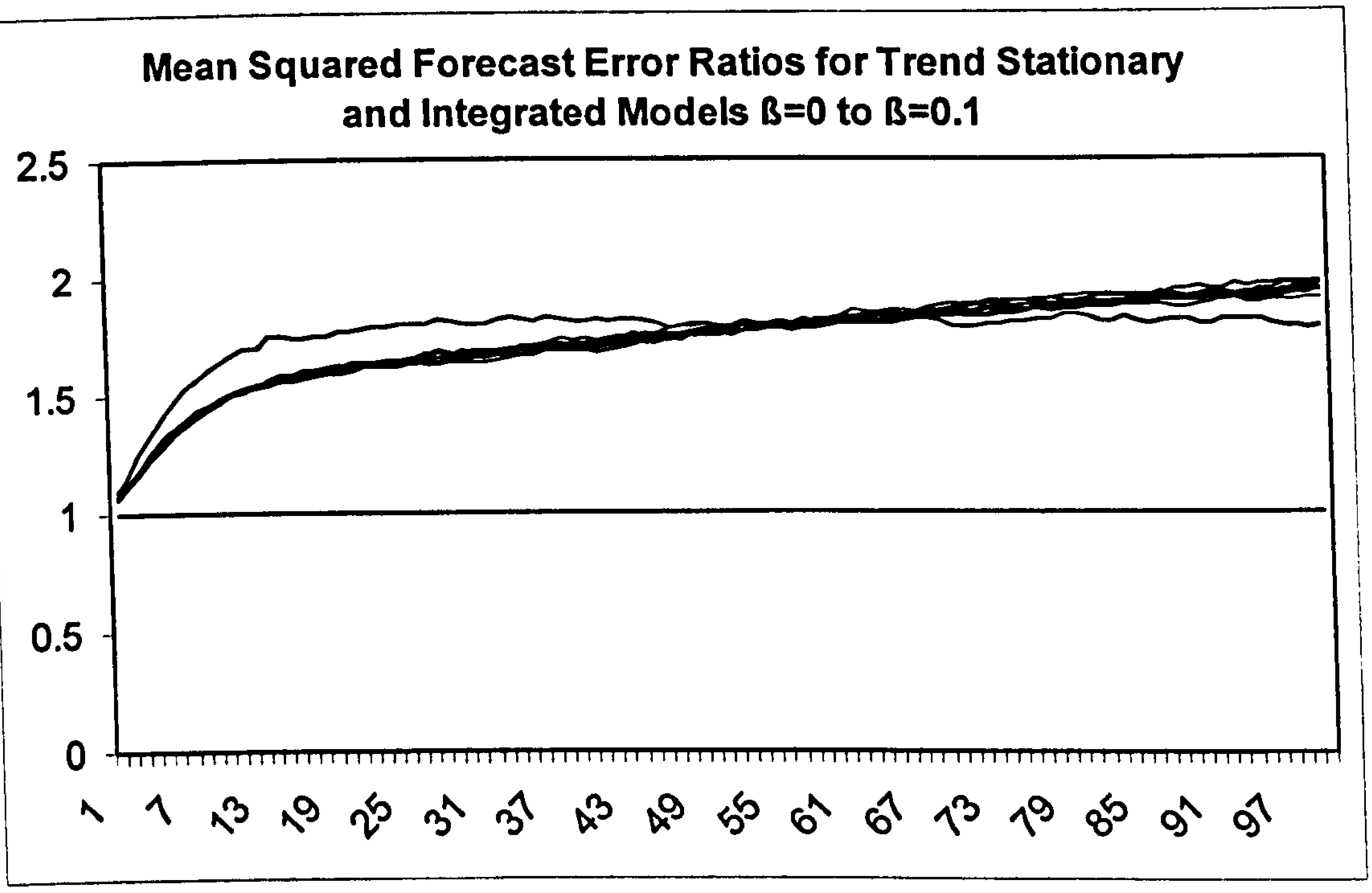
Figure 5.1.9 Simulation Evidence:



DGP is $p_t = a + \beta t + u_t$, $u_t - 0.8u_t = \varepsilon_t$. Number of replications: 20000. Except for the case where $\beta = 0$ mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

2. ***Comparing forecasts from models where the order of integration has been inferred correctly with those where this was not the case.*** Using the same methodology as for $\phi = 0.9$, it is again observed that better forecast results -lower mean squared forecast errors- are obtained when the order of integration of the data generating process is modelled correctly. Again, this holds over the entire 100 period forecast horizon and for all the coefficient values considered. This is illustrated in Figure 5.1.10. below, where the ratio of mean squared prediction errors from the counterfactual model over those from the model corresponding to the DGP consistently takes values above one.

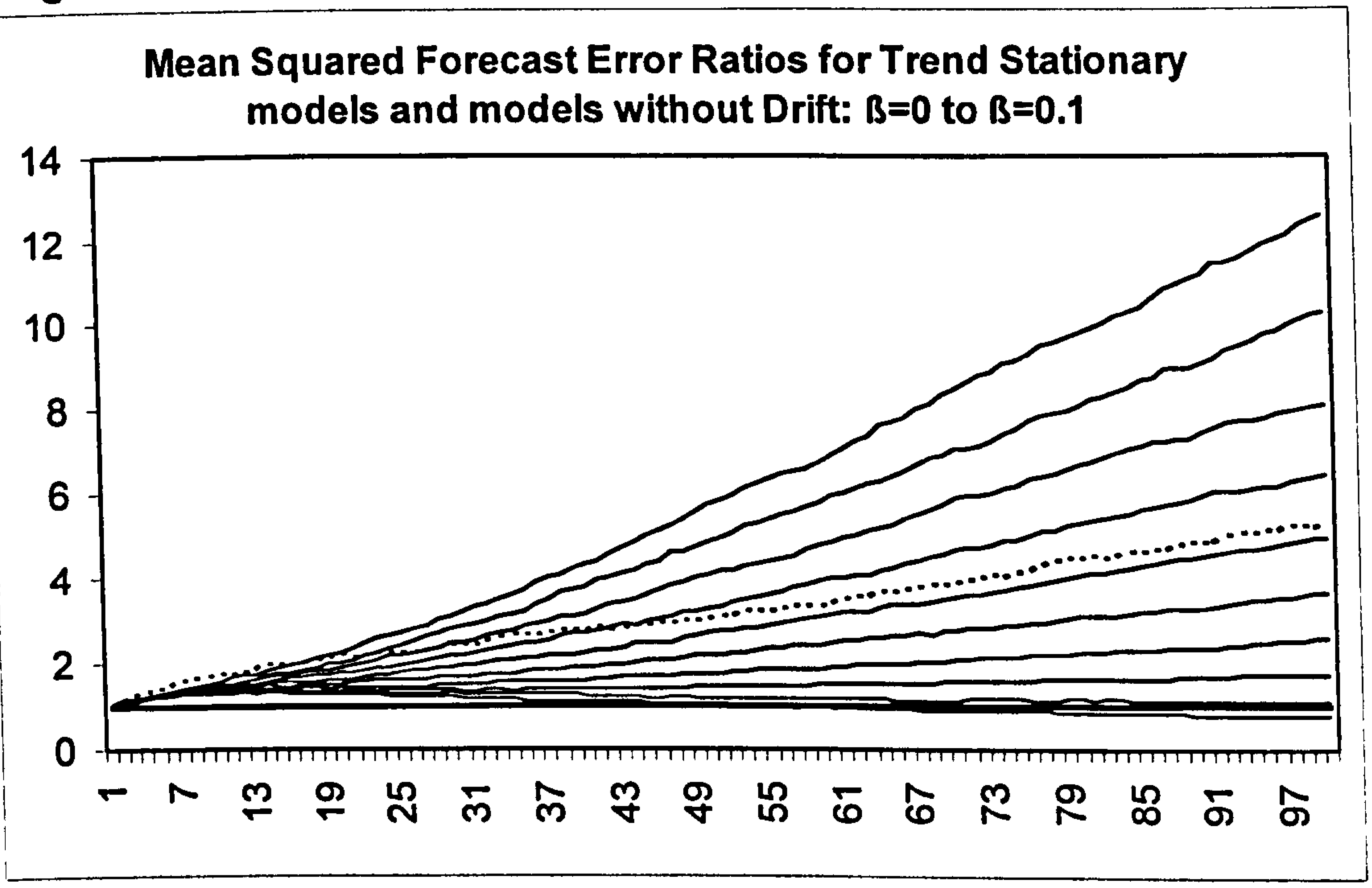
Figure 5.1.10 Simulation Evidence:



DGP is $p_t = a + \beta t + u_t,\ u_t - 0.8u_t = \varepsilon_t$. Number of replications: 20000. Mean Squared Errors for models in levels are in the denominator. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model. For $\beta = 0$ the comparison is between forecasts from models without trend or drift.

3. Considering mistaken inference on both the order of integration and the presence of a trend term. Here again the methods of comparison adopted are the same as for the case where $\phi = 0.9$. Wrongly modelling the trend stationary series as ARIMA(0,1,1) without drift now yields smaller forecast errors only when the trend coefficient in the data generating process takes a value of $\beta = 0.01$ and beyond a forecast horizon length of h=56 or above. For all the other trend coefficient values considered, the model parameterisation corresponding to the data generating process yields smaller mean squared forecast errors over the entire 100 period forecast horizon. The general pattern is illustrated in figure 5.1.11 below.

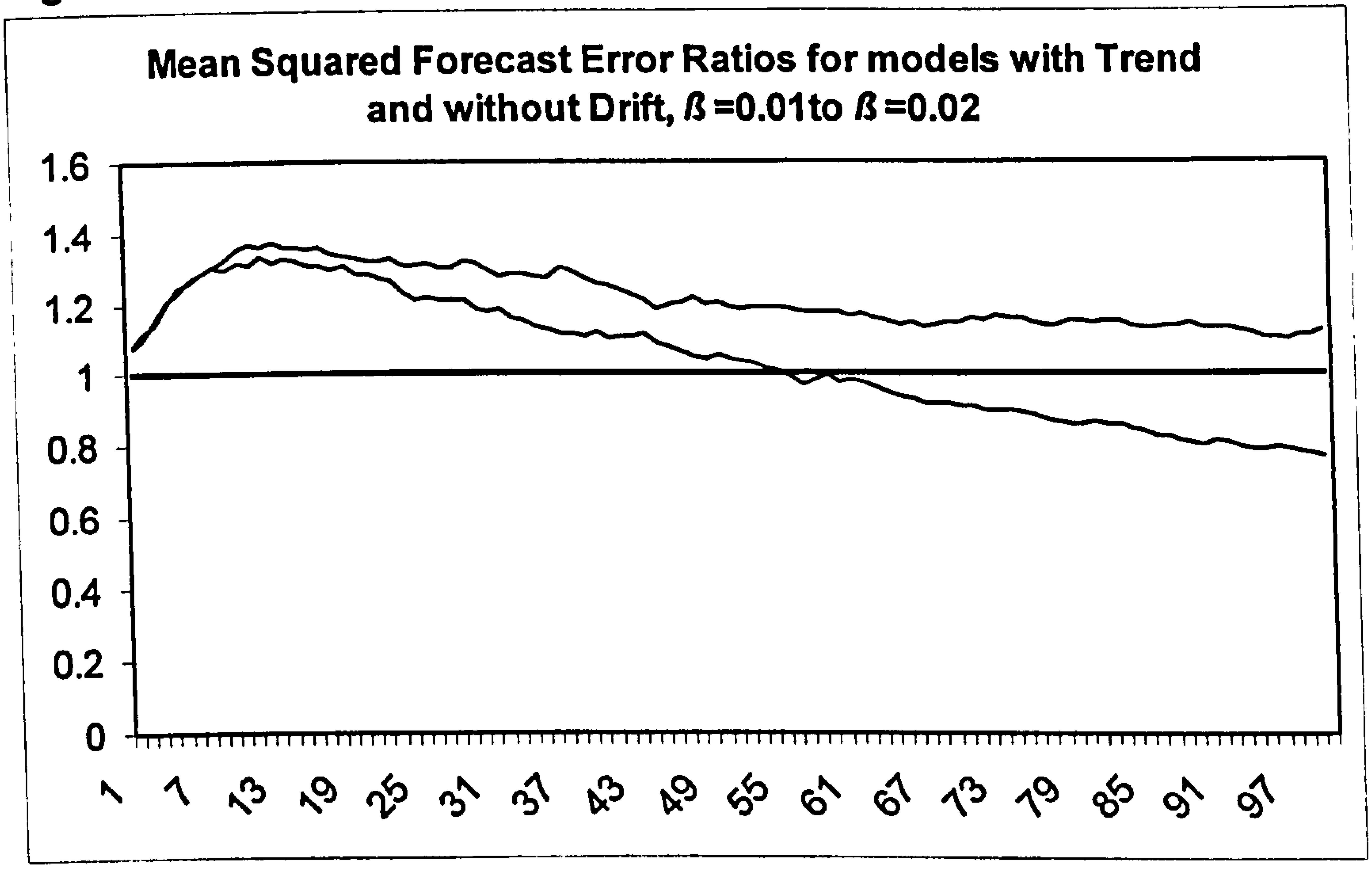
Figure 5.1.11 Simulation Evidence:



DGP is $p_t = a + \beta t + u_t$, $u_t - 0.8u_t = \varepsilon_t$. Number of replications: 20000. Mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

As in Figure 5.1.6 above, the dotted line represents the case of $\beta = 0$, and again, it is apparent that the correct model specification yields superior results for any but very low trend coefficients and at long forecast horizons. This latter case is shown in Figure 5.1.12 below.

Figure 5.1.12 Simulation Evidence:



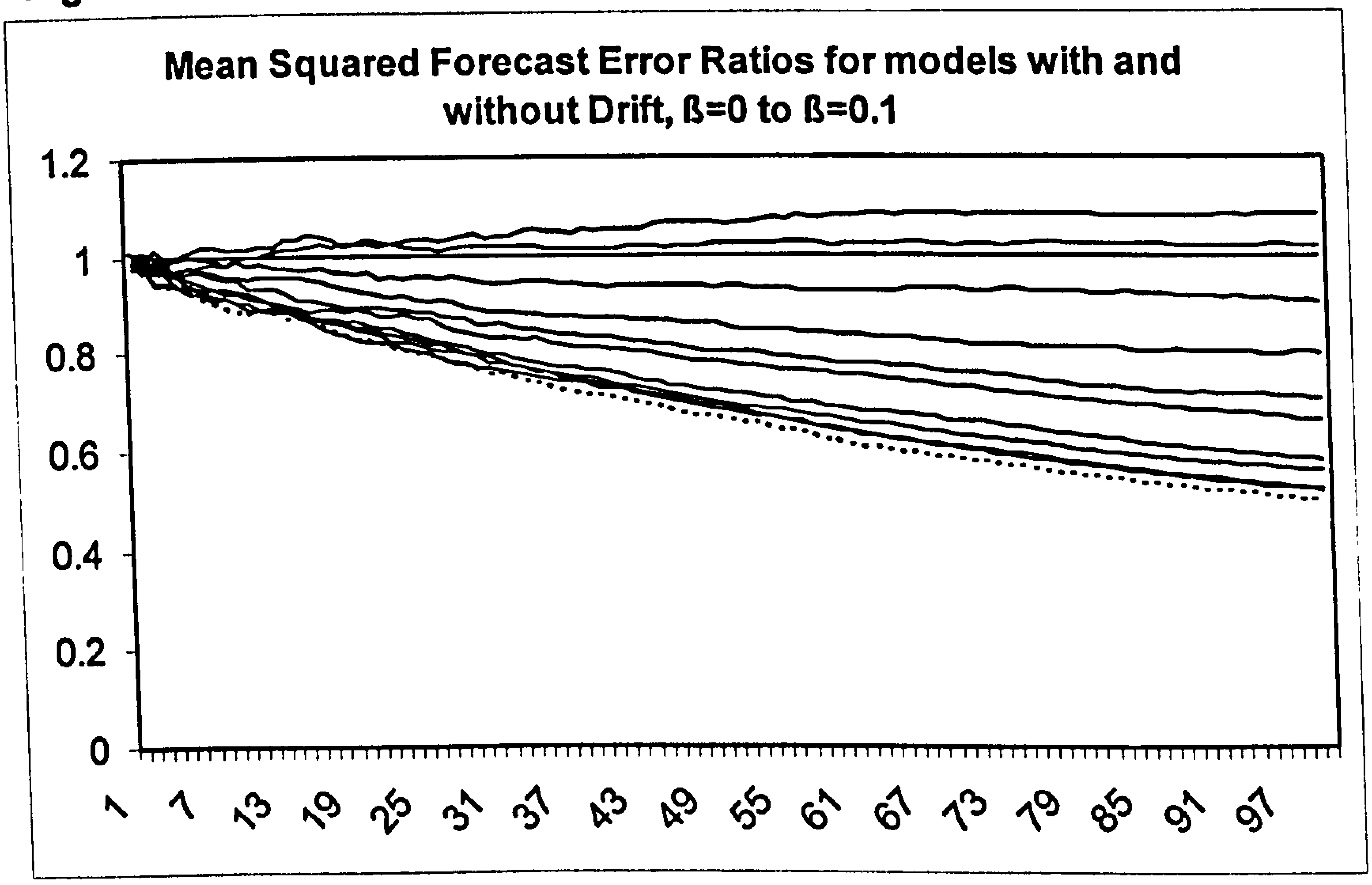
DGP is $p_t = a + \beta t + u_t$, $u_t - 0.8u_t = \varepsilon_t$. Number of replications: 20000. Mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

Figure 5.1.12 shows how the counterfactual forecast model can yield superior results at $\beta = 0.01$ and for $h \geq 56$, while for $\beta \geq 0.02$ consistently superior forecasts are obtained from the correct model specification. (Given the long forecast horizon required for the counterfactual model with $\beta = 0.01$ to yield better results on average this phenomenon is not relevant for model selection in the present case where models are selected for a forecast period of 10 to 20 years only.)

DPG = ARIMA(0,1,1) with $\theta = 0.1$

1. Allowing for erroneous inferences on the presence of a drift term in the data generating process. The statistic employed here is the ratio of the mean squared forecast error for an ARIMA(0,1,1) model without drift divided by the prediction mean squared error for an ARIMA(0,1,1) model including drift, for all data generating processes with a non zero drift term. A ratio above one again indicates that the prediction mean squared error is larger for the counterfactual model. Where the drift term is zero in the DGP, applying this ratio for the assessment of forecast accuracy implies that the counterfactual model now appears in the denominator. In this particular case a value of the mean squared forecast error ratio below one indicates that the true model yields superior forecasts. The general results of this comparison are illustrated in Figure 5.1.13.

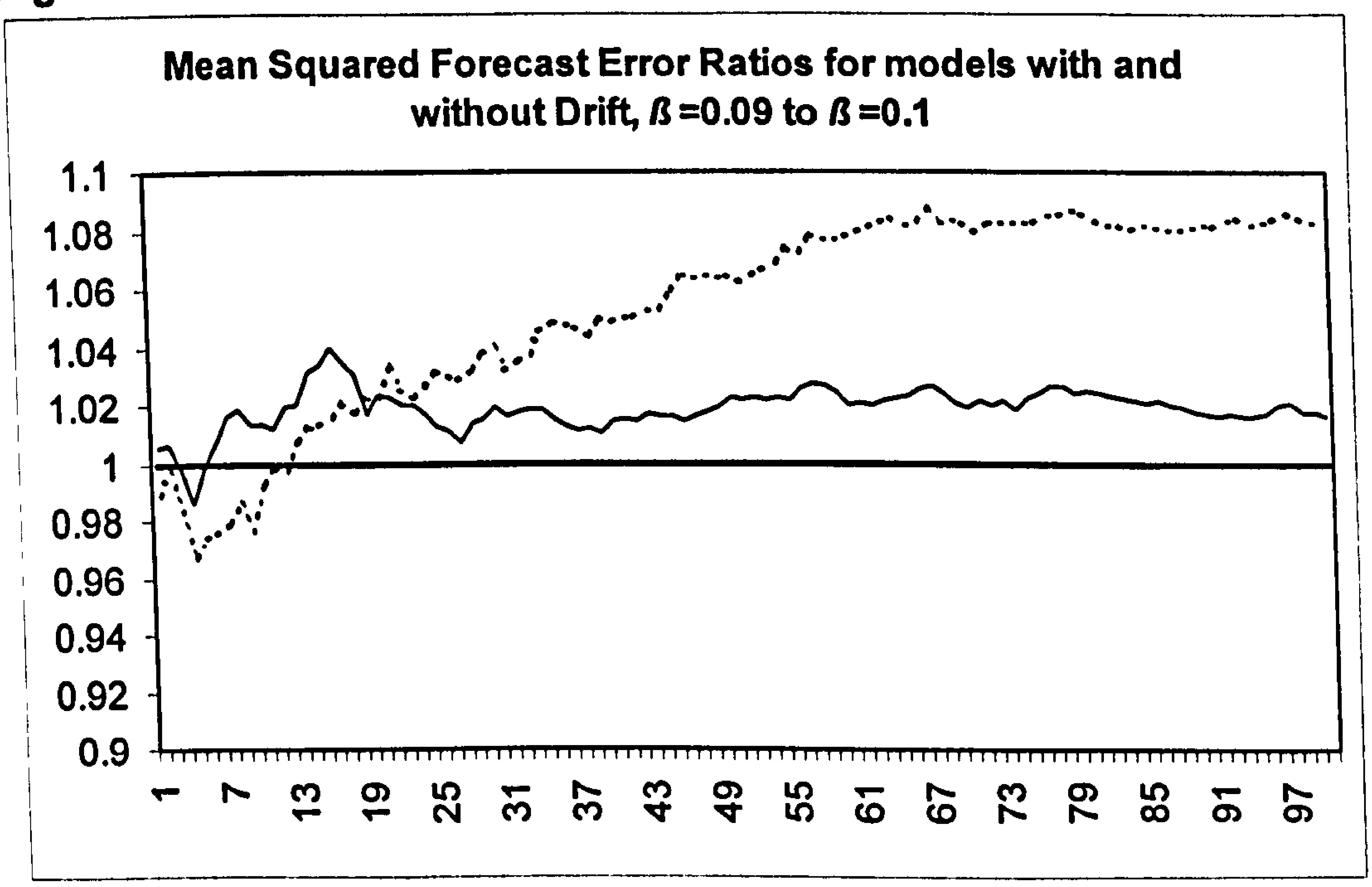
Figure 5.1.13 Simulation Evidence:



DGP is $\Delta p_t = \beta + v_t$, $v_t = \varepsilon_t - 0.1\varepsilon_{t-1}$. Number of replications: 20000. Except for the case where $\beta = 0$ mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the forecast model with trend.

In Figure 5.1.13, a lower position of the line illustrating the mean squared forecast error ratio correspond to a lower drift coefficient value in the data generating process. Where the value of the drift coefficient in the data generating process is zero by construction, the correct model parameterisation yields more accurate forecasts than the corresponding I(1) model with drift (as illustrated by the dotted line in Figure 5.1.13). For non zero drift coefficients below a value of $\beta = 0.09$, forecasts from a model without drift still yield more accurate forecast results on average. As the drift coefficient value reaches 0.09, the ratio of mean squared forecast errors stays close to one. For a drift coefficient value of $\beta = 0.1$, the increased accuracy of the model with drift becomes more pronounced. The ratio of forecast errors takes values clearly above one, although only after 13 forecast periods. The mean squared forecast error ratios for drift coefficients of 0.09 and 0.1 in the data generating process are illustrated in Figure 5.1.14, where the mean squared forecast error ratio corresponding to $\beta = 0.1$ is shown by the dotted line.

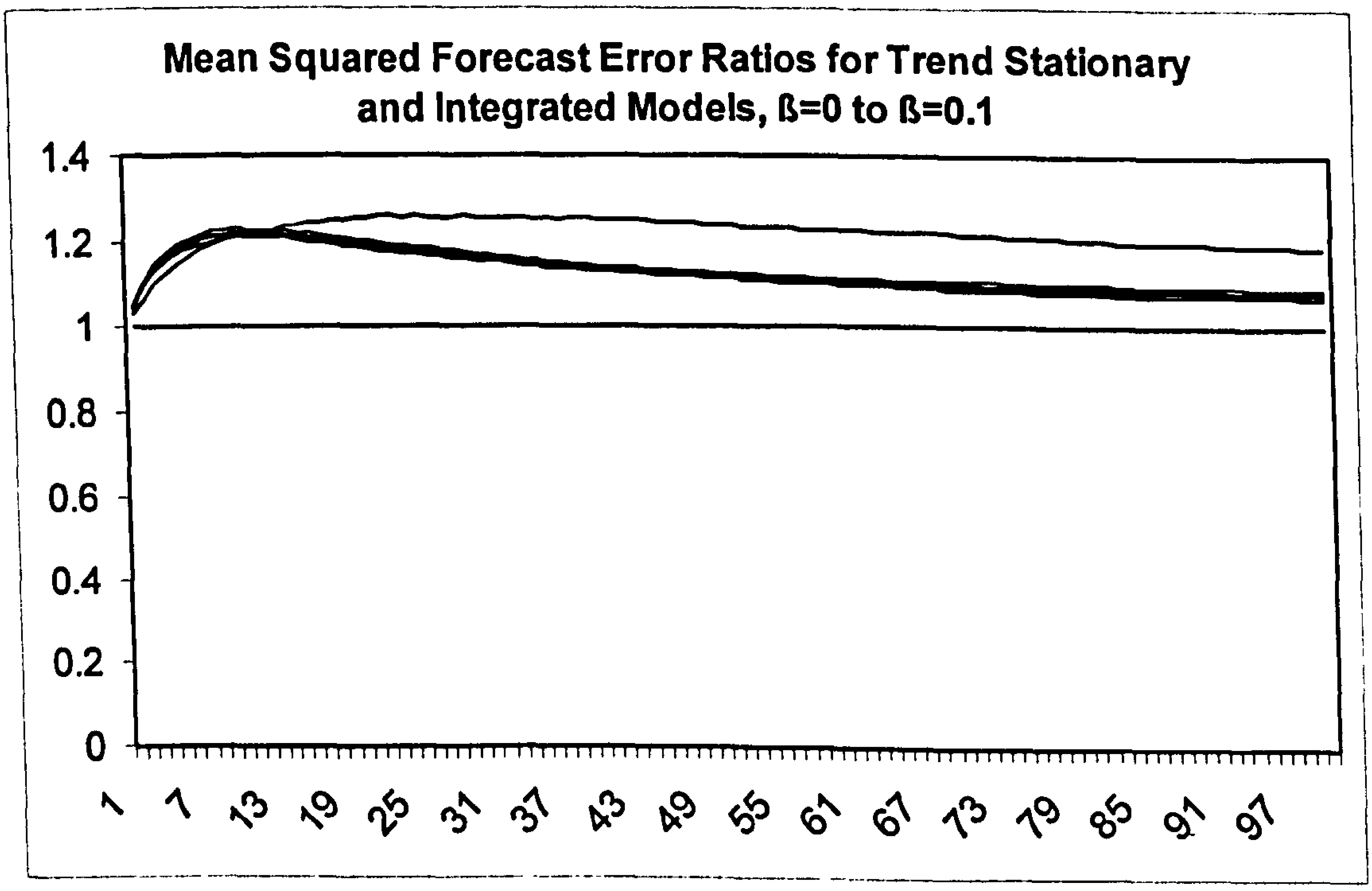
Figure 5.1.14 Simulation Evidence:



DGP is $\Delta p_t = \beta + v_t$, $v_t = \varepsilon_t - 0.1\varepsilon_{t-1}$. Number of replications: 20000. Mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

2. ***Comparing forecasts from models where the order of integration has been inferred correctly with those where this was not the case.*** The comparison made here is between forecast models in levels without trend and forecast models in first differences without drift for the case where the drift coefficient is zero in the data generating process. For all data generating processes with non zero drift coefficients the comparison is between forecast models in levels with trend and forecast models in first differences including drift. It is observed that forecasts from models where the order of integration is specified correctly yield superior forecasts throughout. The extent to which forecasts from the correctly specified model outperform those from a model in levels diminishes at long forecast horizons. This result is shown graphically in Figure 5.1.15 below.

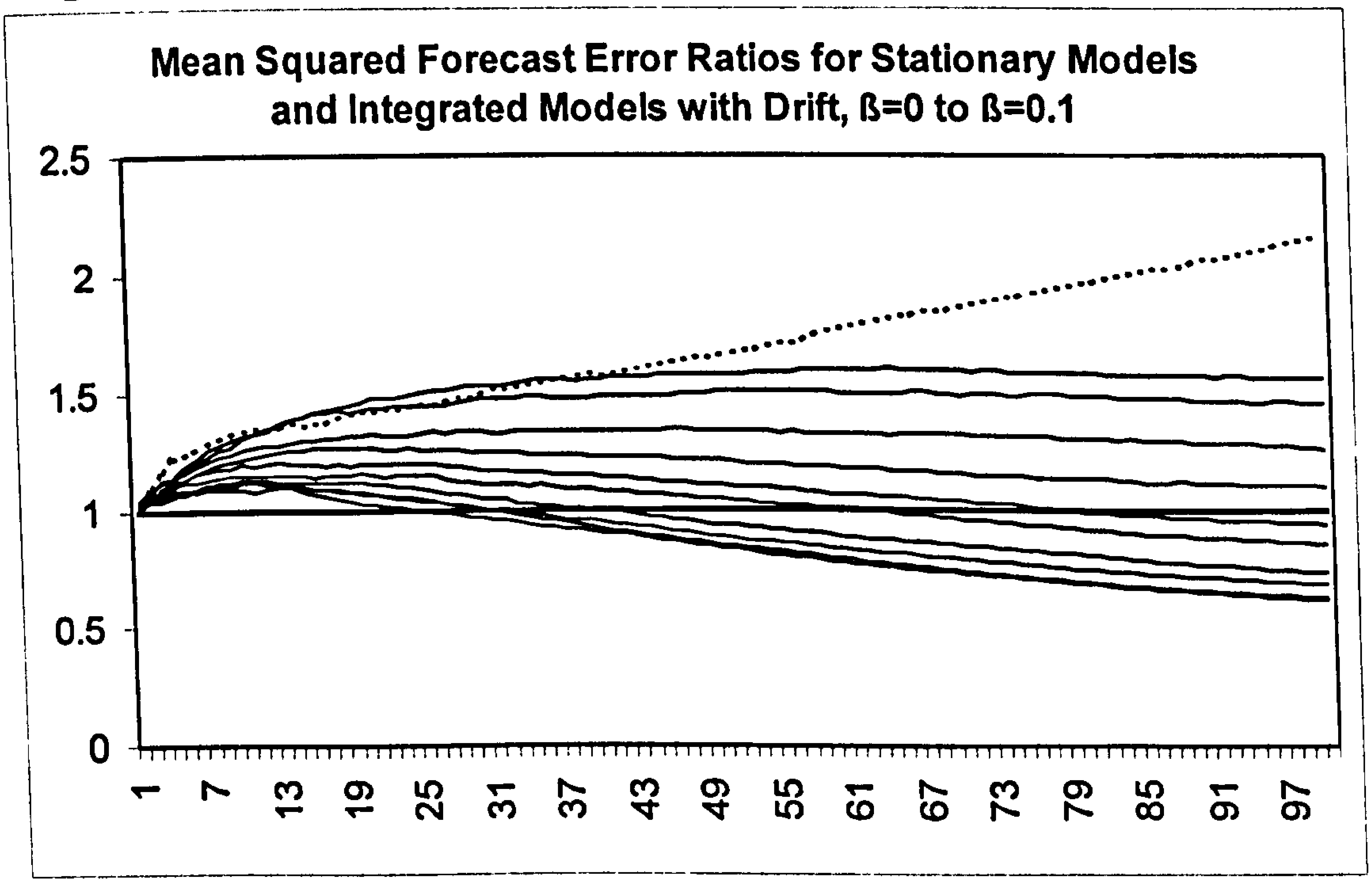
Figure 5.1.15 Simulation Evidence:



DGP is $\Delta p_t = \beta + v_t$, $v_t = \varepsilon_t - 0.1\varepsilon_{t-1}$. Number of replications: 20000. Mean Squared Errors for integrated models are in the denominator. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model. For $\beta = 0$ the comparison is between forecasts from models without trend or drift.

3. Considering mistaken inference on both the order of integration and the presence of a drift term. In the case where the data generating process does not contain a drift component, the comparison is made by taking the ratio of the mean squared forecast errors from a forecast model in levels including a trend to the mean squared forecast error of forecasts from a difference stationary model excluding drift. In the case of a data generating process with a non zero drift component the basis of assessment is the ratio of the mean squared error of forecasts from a model in levels excluding trend over the mean squared forecast error from a difference stationary forecast model with drift. In the driftless case, the correct model specification gives the best forecast results, as illustrated by the dotted line in Figure 5.1.16.

Figure 5.1.16 Simulation Evidence:



DGP is $\Delta p_t = \beta + v_t$, $v_t = \varepsilon_t - 0.1\varepsilon_{t-1}$. Number of replications: 20000. Mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

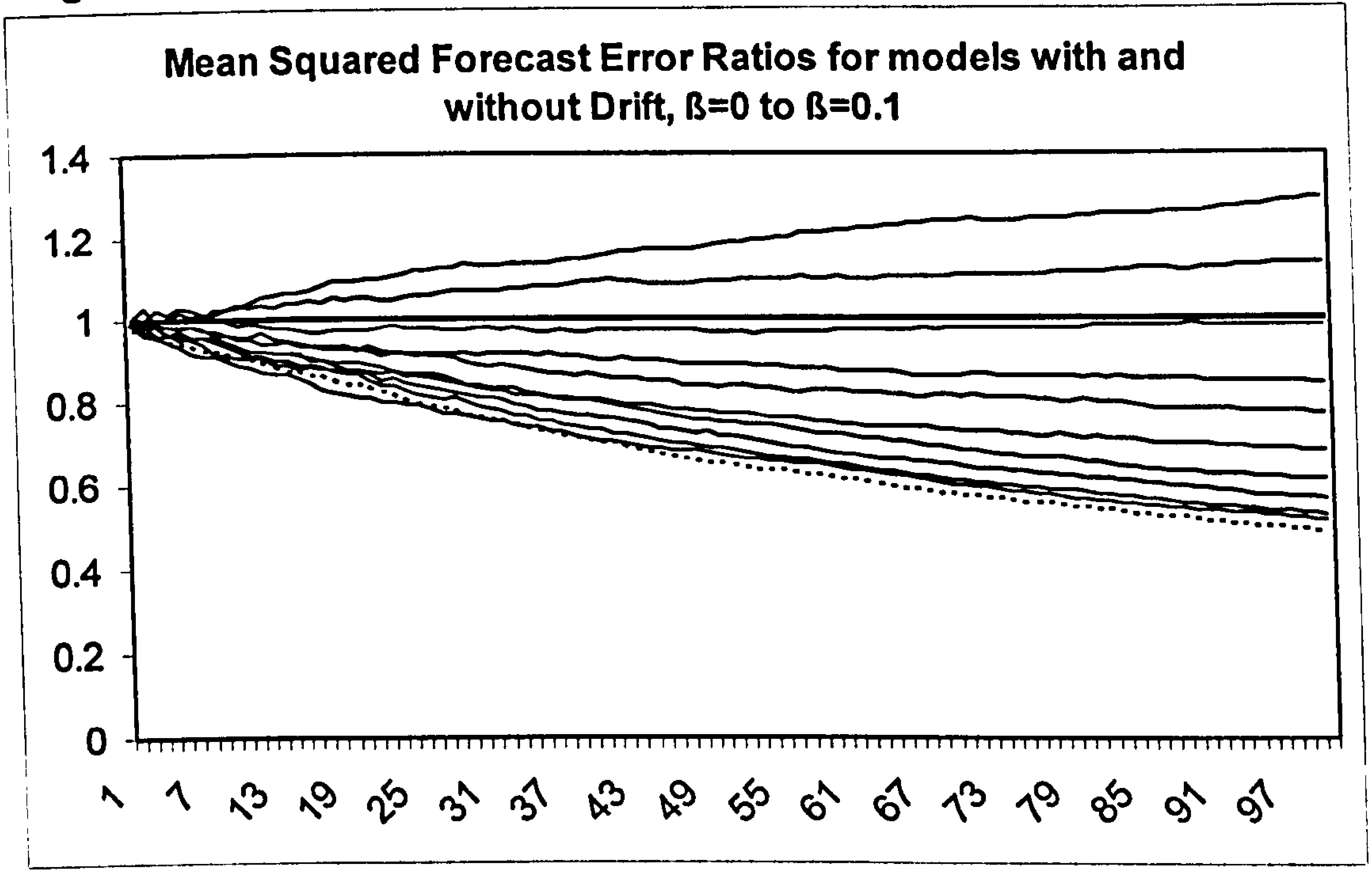
The case of a trend stationary model fitted to a driftless difference stationary process is, of course, of interest in the present study since it can be seen as representative of the case where the presence of a trend term is wrongly inferred for a stationary model when the true DGP is difference stationary. For data generating processes with non-zero drift components, however, better forecast results can be obtained from a trendless model in levels, where the true drift coefficient value is sufficiently small and the forecast horizon is sufficiently large.[6] In Figure 5.1.16, this is illustrated by the solid lines, where lines in a higher position show mean squared forecast error ratios corresponding to larger drift coefficient values in the data generating process.

ARIMA(0,1,1) with $\theta = 0.2$

1. Allowing for erroneous inferences on the presence of a drift term in the data generating process. The comparison of mean squared error ratios was as for $\theta = 0.1$ above. Again, the correct model specification yields superior forecast results for a data generating process without drift. It is also the case that driftless forecast models outperform I(1) forecast models with drift for low values of the data generating process. The general pattern is illustrated in Figure 5.1.17 below, where the dotted line shows the case of a zero drift coefficient in the DGP and generally, lower positioned lines correspond to lower drift coefficient values.

[6] More precisely, the correct model parameterisation yields better forecast results initially, but then is outperformed by forecasts from the counterfactual forecast model at a time horizon of h=21 forecast periods where the true drift coefficient value is $\beta = 0.01$. The forecast horizon required for this phenomenon to occur then lengthens to h=31 for $\beta = 0.02$ and $\beta = 0.03$, h=41 for $\beta = 0.04$, h=63 for $\beta = 0.05$ and finally h=79 for $\beta = 0.06$. At drift coefficient values of $\beta = 0.07$ and above, the correct forecast model specification yields consistently superior forecasts.

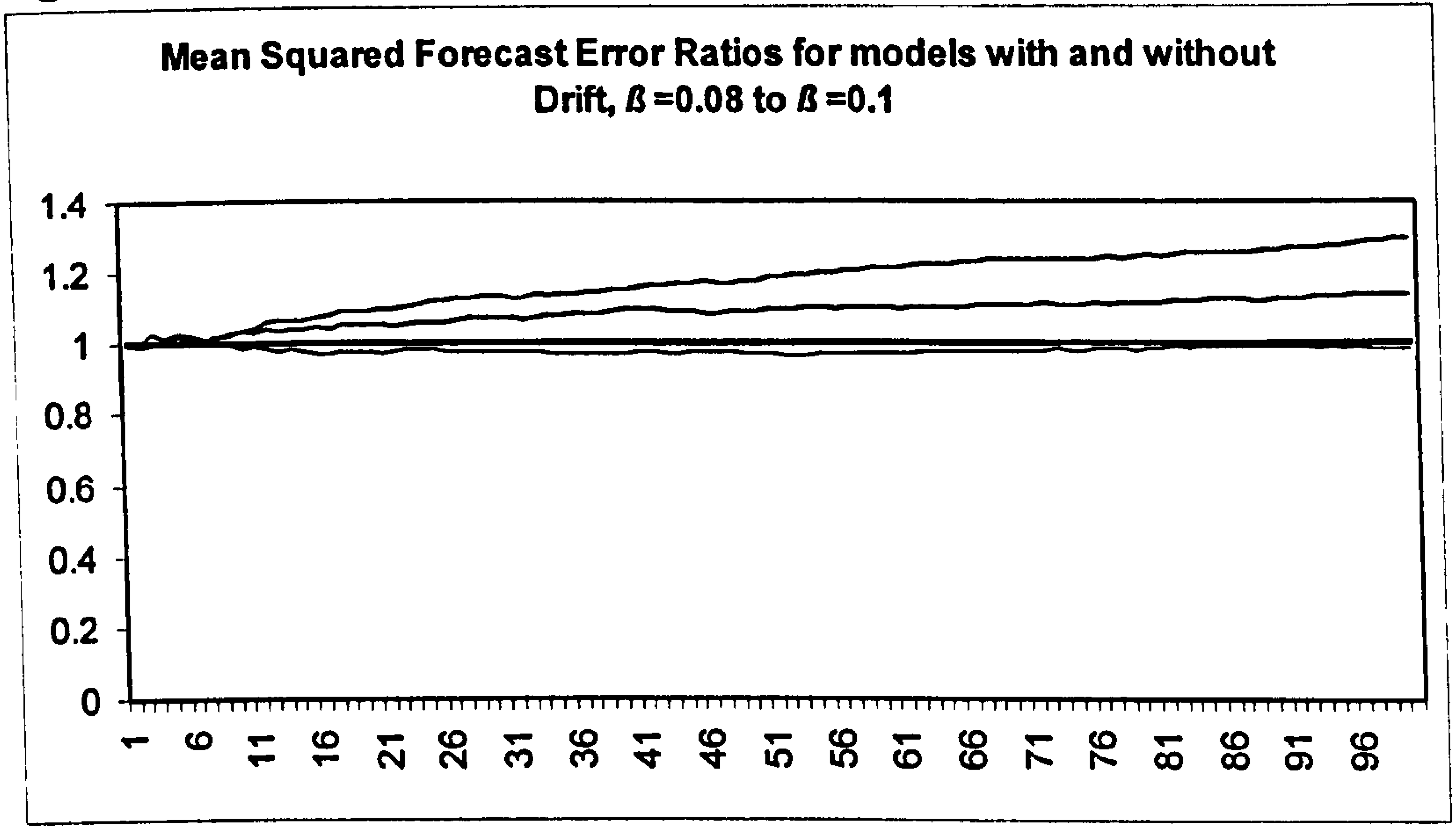
Figure 5.1.17 Simulation Evidence:



DGP is $\Delta p_t = \beta + v_t$, $v_t = \varepsilon_t - 0.2\varepsilon_{t-1}$. Number of replications: 20000. Except for the case where $\beta = 0$ mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

The ratio of mean squared forecast errors converges to values close to one at short forecast horizons for a drift coefficient of $\beta = 0.08$. Even when $\beta = 0.07$ the ratio of mean squared forecast errors stays close to one at very short forecast horizons but then falls to values of 0.95 and below. As the drift coefficient of the DGP reaches values of 0.09 and above, the superior forecast performance of the forecast model with drift becomes apparent. For $\beta = 0.09$, and $h \leq 7$ driftless forecast models still yield prediction mean squared errors which are very close to those from a difference stationary model with drift. At longer forecast horizons, the ratio of prediction mean squared errors rises substantially above one. The case for drift coefficient values between $\beta = 0.08$ and $\beta = 0.1$ is shown below in Figure 5.1.18.

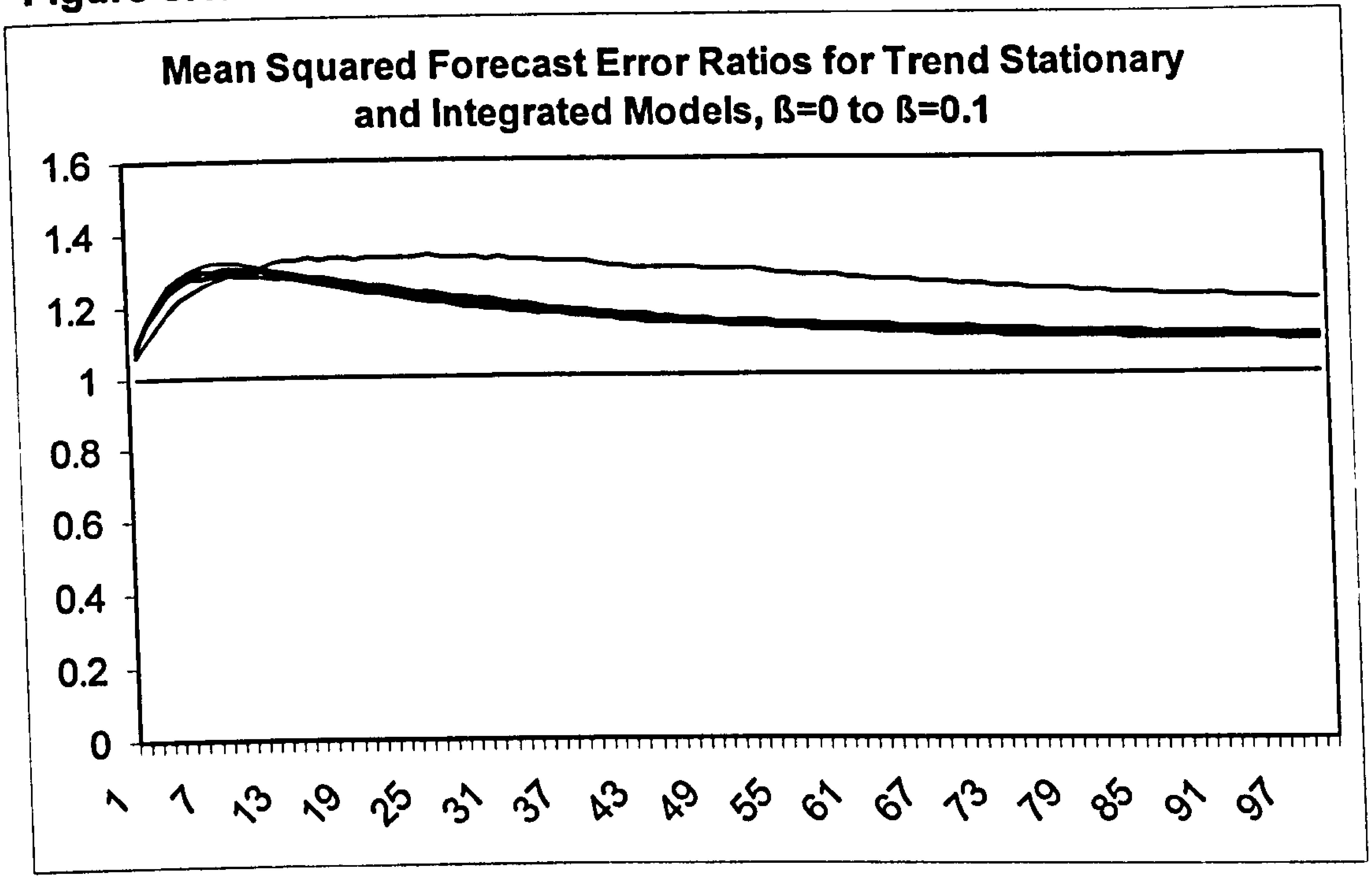
Figure 5.1.18 Simulation Evidence:



DGP is $\Delta p_t = \beta + v_t$, $v_t = \varepsilon_t - 0.2\varepsilon_{t-1}$. Number of replications: 20000. Mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model. Lines in higher positions correspond to larger values for the drift coefficient in the DGP.

2. ***Comparing forecasts from models where the order of integration has been inferred correctly with those where this was not the case.*** Using mean squared forecast error ratios as above for $\theta = 0.1$, it again appears that forecasts from I(1) models consistently outperform forecasts from stationary models where the data generating process is difference stationary. This holds regardless of the value of the drift coefficient, and the extent to which the difference stationary model yields superior forecasts diminishes over long forecast horizons. This result is illustrated in Figure 5.1.19.

Figure 5.1.19 Simulation Evidence:



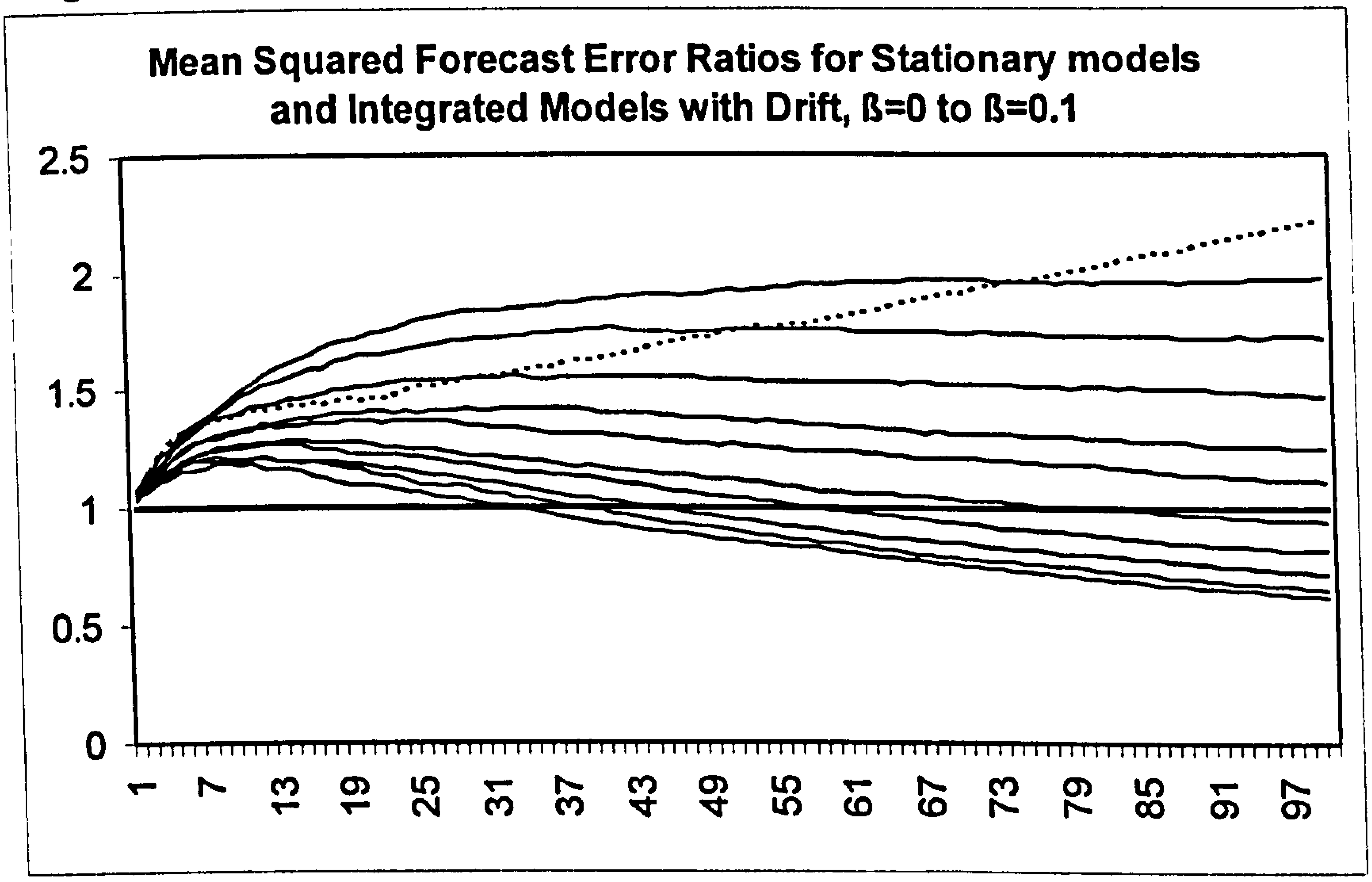
DGP is $\Delta p_t = \beta + v_t$, $v_t = \varepsilon_t - 0.2\varepsilon_{t-1}$. Number of replications: 20000. Mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

3. Considering mistaken inference on both the order of integration and the presence of a drift term. The possibility of misspecifying both the order of integration and the presence of a drift term was again tested as above. It was again observed that the correct model specification yields superior forecasts when the data generating process does not contain a drift term (*i.e.* in the case relevant for spurious rejections of the zero trend null hypothesis). For drift coefficient values between 0.01 and 0.05, it is observed that misspecified models can yield superior forecasts at long forecast horizons[7].

[7] In more detail, the misspecified model here yields superior forecasts beyond h=33 when $\beta = 0.01$, h=39 for $\beta = 0.02$, h=45 for $\beta = 0.03$, h=58 for $\beta = 0.04$ and h=82 in the case of $\beta = 0.05$. For drift coefficients of $\beta = 0.06$ and above, the correctly specified model yields superior forecasts over the entire horizon considered.

The results are shown in Figure 5.1.20 below. The dotted line illustrates the case for $\beta = 0$ in the DGP, where the mean squared forecast error is consistently larger for the counterfactual model. The solid lines represent the mean squared forecast error results for data generating processes with non zero drift coefficients. Lines in a lower position are representative of lower drift coefficient values in the DGP.

Figure 5.1.20 Simulation Evidence:



DGP is $\Delta p_t = \beta + v_t$, $v_t = \varepsilon_t - 0.2\varepsilon_{t-1}$. Number of replications: 20000. Mean Squared Errors for the correct forecast model specification are in the denominator throughout. The Forecast horizon is denoted on the X-axis. Values larger than one indicate larger Mean Squared Errors for the counterfactual forecast model.

Some of the pattern of relative forecast performance that emerge from the simulation results above are what one would expect. Correctly identifying the order of integration of the data generating process has been shown to consistently improve forecast performance. Superior forecasts are also generally obtained when a trend or drift coefficient is rightly excluded. Correctly including a trend or drift component only leads to better forecast results than excluding it if the trend or drift coefficient value is sufficiently large and the forecast horizon sufficiently long. This property appears to be more pronounced for difference stationary than for trend stationary models.

Table 5.1.1 shows the number of forecast periods after which lower forecast errors are obtained when a trend or drift term is included in the forecast model. This minimum forecast horizon is specified for a number of values of β as well as for the trend and difference stationary data generating processes considered here. (Note that the order of integration of the forecast model used corresponds to the order of integration of the DGP. The underlying comparisons are between models with and without trend or drift component only.)

Table 5.1.1 Forecast horizon after which forecasts from models with trend or drift have lower forecast errors than forecasts from models without trend or drift.

β	$\phi = 0.9$	$\phi = 0.8$	$\theta = 0.1$	$\theta = 0.2$
0	-	-	-	-
0.01	-	-	-	-
0.02	-	0	-	-
0.03	4	0	-	-
0.04	0	0	-	-
0.05	0	0	-	-
0.06	0	0	-	-
0.07	0	0	-	-
0.08	0	0	-	-
0.09	0	0	6	4
0.1	0	0	13	0

β: Value of the trend/drift coefficient in the DGP, ϕ: AR(1) coefficient where the DGP is ARIMA(1,0,0),
θ: MA(1) coefficient where the DGP is ARIMA(0,1,1), '-': The number of forecast periods required exceeds 100.

It can moreover be observed that the relative increase in the mean squared error resulting from mistaken inference about the presence of a trend coefficient is larger than the impact of mistaken assumptions about the presence of a drift term in the forecast model. Somewhat counterintuitive results suggesting possible improvements from forecast models based on misspecifications of both the order of integration and the presence of trend components are practically irrelevant here since they are confined to longer forecast horizons than are considered in the present case.

In a final comparison of mean squared forecast errors, an attempt is made to assess the relative cost of fitting the wrong model. This is clearly of relevance where substantial uncertainty about model selection can not be overcome. For this purpose the ratio of mean squared forecast errors for forecasts from a trend stationary model fitted to a difference stationary data generating process without drift can be compared with the mean squared forecast error for forecasts from a difference stationary model without drift fitted to a trend stationary process[8]. Such a comparison indicates that the forecast errors from a wrongly fitted difference stationary model without drift are consistently below those from a wrongly fitted trend stationary model. The only exception to this occurs at forecast horizons of up to two periods where the ratio of forecast errors from a wrongly fitted trend stationary model to those from a wrongly fitted model in first differences fall to values just below one.

[8] The decision to wrongly include a trend term in the stationary model while omitting a drift term from the I(1) model fitted to a trend stationary DGP is obviously motivated by the fact that these mistaken inferences are the ones that are likely to occur where conclusions on the presence of a trend component in the data series depend on assumptions regarding the order of integration.

It has been shown so far that the inferred order of integration as well as inference on the trend or drift term have an important influence on the quality of the forecasts obtained. It had been shown in chapters three and four above that both inferential tasks can be persistently interrelated. In the following, it will therefore be investigated if and how Diebold and Killian's conclusions change when different magnitudes of the trend coefficient estimate and interdepence of the order of integration and the detection of a trend are incorporated into the simulation experiment.

5.1.5 Diebold and Killian Revisited

The above observations on the relative forecast performance of models with and without trend make it seem advisable to look further into the results on the use of pre-testing reported by Diebold and Killian (2000). The overall results on pre-testing prove to be fairly robust to variations in the size of the trend coefficient. The simulation experiment for the replication of the Diebold and Killian results -as described above- was here conducted for a number of values for the autoregressive coefficient and the trend coefficient again using up to 100 forecast periods and 20,000 replications. The trend coefficient for the data generating process (*i.e.* β) was allowed to take values of between zero and 0.13, increasing the coefficient values by 0.01 at a time. (The coefficient value imposed in the Diebold and Killian experiments is $\beta = 0.0065$, *or* $\beta/\sigma_{\varepsilon} = 0.65$ by comparison.) The values used for the autoregressive coefficient are 0.8, 0.85, 0.9, 0.95, 0.97, 0.99 and 1. The difference stationary alternative to the AR(1) model with trend was a random walk with drift in all cases.

Taking the ratio of prediction mean squared errors for the differenced alternative model specification over forecasts from the model selected by pre-testing yields very similar results across all trend coefficient values (including $\beta = 0$). For data generating processes with autoregressive parameters of phi=0.8 and phi=0.85 forecasts from the pre-test model outperform those from a continually differenced model over all time horizons and over all trend coefficient values. As the autoregressive coefficient rises to phi=0.9 this is generally true after the first five forecast periods[9]. For an AR coefficient value of phi=0.95, the ratio of prediction mean squared errors converges to one over all trend coefficient values. As the value of the autoregressive coefficient rises above phi=0.97 the differenced model consistently outperforms the model selected by pretesting, *i.e.* the prediction mean squared error ratio starts dropping below one over the entire forecast horizon for all trend coefficient values considered.

These results are generally in line with the replication of Diebold and Killian's results described above. The results therefore seem to be robust to the choice of different trend coefficient values. However, one particularly relevant question in the present case is that of fitting either a trend stationary model or a difference stationary model without drift. These are the alternatives that would tend to arise if one relied entirely on pre-testing and the use of standard t-tests on the trend coefficient when choosing the forecast model.

To investigate the impact of using pretests for forecast model selection when the alternatives considered are an ARIMA(1,0,0) model including trend and constant

---

[9] For $\beta = 0.01$ this is true after the first four forecast periods

and a difference stationary model without drift a new set of simulation experiments was conducted. Simulations were again conducted as above in section 5.1.2, *i.e.* with samples of size T=100, a forecast horizon of up to h=100 and 20,000 replications. Again the trend coefficients for the data generating process took the following values: 0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.11, 0.12 and 0.13. The values considered for the autoregressive coefficient were again as above with phi equal to 0.8, 0.85, 0.9, 0.95, 0.97, 0.99 or 1. To allow comparisons with the results by Diebold and Killian as well a with the more general simulation results reported above, the forecast model alternatives considered were: ARIMA(0,1,0) without drift, *i.e.* a random walk, and ARIMA(0,1,1) without drift.

Contemplating the values of prediction mean squared error ratios over forecast periods (h) of up to h=100, it can be seen that for both, the ARIMA (0,1,1) and for the random walk alternatives, the differenced model consistently outperforms the model selected by pretesting for autoregressive coefficient values of phi= 1, 0.99 and 0.97 as before. For $\phi = 0.95$, the differenced model dominates[10] the pre-test model for trend coefficient values within the range $0 \leq \beta \leq 0.1$. As the trend coefficient rises beyond 0.1, the pre-test model outperforms the differenced model after 83, 59 and 41 forecast periods for trend coefficient values of 0.11, 0.12 and 0.13 respectively when the difference stationary model alternative is ARIMA (0,1,1). For the case of a random walk, the pre-test model outperforms the

[10] *I.e.* it has lower mean squared prediction errors than the pre-test model.

differenced model after 84, 58 and 41 periods for trend coefficient values of 0.11, 0.12 and 0.13 respectively.

As the value of the autoregressive coefficient falls to $\phi = 0.9$, the difference stationary alternative yields superior forecasts over the entire forecast horizon for trend coefficient values of $0 \leq \beta \leq 0.6$. Thereafter, the pre-test model performs better after a forecast horizon of 56, 13, 10, 8, 7, 7, and 6 periods for trend coefficient values of 0.07, 0.08, 0.09, 0.1, 0.11, 0.12 and 0.13 respectively in the case of the ARIMA (0,1,1) specification for the difference stationary model. For the random walk case, the pre-test model achieves lower forecast errors after forecast periods of length 56, 11, 9, 9, 7, 7 and 6 for trend coefficient values of 0.07, 0.08, 0.09, 0.1, 0.11, 0.12 and 0.13 respectively.

At an autoregressive coefficient value of $\phi = 0.85$, the random walk model yields superior results for $0 \leq \beta \leq 0.04$. For the ARIMA (0,1,1) model, the pre-test model has smaller forecast errors on average after two forecast periods for trend coefficient values of $\beta = 0.05$ and $\beta = 0.06$. For higher trend coefficient values, the pre-test model dominates over the entire forecast horizon. In the case of the random walk alternative, forecasts from pre-test models have smaller prediction mean squared errors after one period for all values of $\beta > 0.04$ considered in the simulation experiment.

At a value of $\phi = 0.8$ finally, the ARIMA (0,1,1) model outperforms the model selected by pre-testing for trend coefficient values of $\beta = 0$, $\beta = 0.01$ and $\beta = 0.02$ over the entire 100 period forecast horizon. For all other trend coefficient values used in the simulations with values of $\beta \geq 0.03$ the model selected by pre-testing performs better than the consistently applied differenced model over the entire forecast horizon. When a random walk is considered as a difference stationary alternative model, the model selected by pre-testing yields smaller forecast errors for all trend coefficients considered, except $\beta = 0$ and $\beta = 0.01$.

It can be concluded that pre-testing for unit roots can in some cases improve the forecast performance of a univariate time series model. As before, such an improvement is unlikely at values for the autoregressive coefficient close to one. In contrast to the results presented initially though, different conclusions can be reached for lower values of the autoregressive coefficient if the alternative model specification is a difference stationary model without drift and the trend coefficient in the data generating process is sufficiently small. In this case, the consistent application of differenced models can yield superior forecast results at small trend coefficient values as well as at high values for the autoregressive coefficient.

5.2. Selecting forecast models for different commodities

It was argued above that the order of integration of the series as well as the trend or drift coefficient are crucial in determining the adequate forecast model. (Of course this is also true of the forecast model more generally and the value of the AR(1) coefficient in particular.) The following general points should be emphasised:

- The importance of the order of integration -compared to the presence of absence of a trend or drift component- in the data series. Consideration of the correct order of integration is important because of the direct influence on the forecasts obtained as well as in those cases where it is difficult to reach conclusions on the significance of the trend or drift component independently of conclusions on the order of integration of the series.

- Aside from the presence of a trend or drift coefficient in the data generating process, the suitability of using forecast models with trend or drift can be shown to depend on the magnitude of the trend coefficient. (This is of particular importance for the prediction of integrated processes.) If there is sustained uncertainty about the appropriate model specification and the drift coefficient values under consideration take sufficiently low values, the smaller cost of fitting the wrong model is likely to arise from fitting a difference stationary model without drift.
- There can be some improvements in forecast performance from pre-testing. At very low trend coefficient values and for near integrated series fitting a differenced model without trend can still be the superior solution.

Following these criteria, the selection of forecast models for individual commodity price series is made below. It should be remembered though that one should be careful in generalising from the simulation results reported above. Results differ for the different model parameterisations considered in the simulation experiments and further differences are to be expected for yet different model parameterisations and at different coefficient values. One cannot therefore generally apply the above conclusions unless the estimated model alternatives are very close to those considered in the simulations.

Moreover, where significant drift coefficients are only obtained when selecting the difference stationary model by AIC, this criterion will be employed to select the forecast model. Where the optimal forecast model identified in the difference stationary case and on the basis of the SBC is a pure random walk or a random walk with drift, an ARIMA(1,1,1) specification will be considered alternatively.

To facilitate a general comparison with the above simulation results, trend and drift coefficient estimates for the model alternatives discussed below were normalised with respect to the estimated standard error of the residual of the ARIMA model in question, *i.e.* where reference is made to normalised coefficients the coefficient is expressed as:

[5.2.1] $$\tilde{\beta} = \frac{\hat{\beta}}{\hat{\sigma}_\varepsilon},$$

where $\tilde{\beta}$ is the normalised trend or drift coefficient, $\hat{\beta}$ is the original trend or drift coefficient estimate and $\hat{\sigma}_\varepsilon$ is the standard error of the residual in the estimated ARIMA model[11].

5.2.1. Selected Forecast Models for Individual Price Series

Focusing initially on those models where the presence of a trend was inferred with a high degree of confidence, the selected forecast models are as follows:

Aluminium: This is identified as stationary in levels by the ADF unit root test as well as the Leybourne McCabe test. This stationarity conclusion is strengthened by the parameter estimates obtained for the trend stationary and differenced model. The point estimates for the trend stationary model are:

[5.2.2] $$p_t = 1.356 - 0.019t + u_t,\ u_t - 0.650u_{t-1} = \varepsilon_t + 0.472\varepsilon_{t-1}$$

Differencing this expression and defining $(1-L)u_t = v_t$ yields:

---

[11] Of course, for the computation of the actual forecasts, the non normalised value of the trend coefficient estimate will be used throughout.

[5.2.3] $(1-L)p_t = -0.019 + v_t,\ v_t = 0.650v_{t-1} + 0.472\varepsilon_{t-1} - 0.472\varepsilon_{t-2} - \varepsilon_{t-1} + \varepsilon_t$

or:

[5.2.4] $$\Delta p_t = -0.019 + v_t,\ v_t - 0.650v_{t-1} = \varepsilon_t - 0.528\varepsilon_{t-1} - 0.472\varepsilon_{t-2}$$

The estimated ARIMA(1,1,2) model for Aluminium is:

[5.2.5] $$\Delta p_t = -0.019 + v_t,\ v_t - 0.679 = \varepsilon_t - 0.537\varepsilon_{t-1} - 0.463\varepsilon_{t-2}$$

The estimated difference stationary model selected by minimum SBC not only shows the parameterisation corresponding to an overdifferenced trend stationary model, it also has parameter estimates with values very close to those of the differenced trend stationary model. More importantly, the estimated moving average parameters sum to one in the case of the estimated ARIMA(1,1,2) model. The trend component is shown to be significant independently of stationarity assumptions with normalised point estimates of -0.124 and -0.126 for the trend and drift coefficient respectively. Although the Vogelsang test only provided limited support for the presence of a trend and the Sun and Pantula corrections for serial correlation are not applicable here, the high value of the t-ratios (-8.997 for the trend and -7.931 for the drift coefficient) make it appear likely that coefficient estimates would still appear significant if the finite sample impact of serial correlation had been quantified for the given model parameterisation. Given the magnitude of the point estimates for trend and drift, it is also unlikely that the forecast performance will be adversely affected by inclusion of the trend, even though the above simulation results can not strictly be generalised. Forecasts are therefore made from an I(0) model with trend.

Rubber: Again, the trend and drift coefficients appear statistically significant with normalised point estimates of -0.104 and -0.105 respectively[12] when the difference stationary model is selected by AIC. (It will be recalled from chapter 4 that the trend coefficient still appears significant if the corrections suggested by Sun and Pantula (1999) are implemented and the price series for Rubber is modelled as trend stationary. The result no longer holds though if Sun and Pantula's modified pre-test method is employed.) The ADF test just fails to reject the null hypothesis of a unit root, while the Leybourne McCabe test identifies the series as trend stationary. Given the magnitude of the trend or drift coefficient estimates and an estimated value of 0.796 for the autoregressive coefficient in the ARIMA(1,0,0) model the conclusion that would follow from the above simulation results is such that pre-testing can be usefully employed to improve forecast performance. On the basis of pre-testing alone, the model selection would most likely be made in favour of a difference stationary model[13] if the ADF test is used, although the Leybourne McCabe test fails to reject the stationarity null hypothesis. However, in view of the fact that the moving average coefficients in the difference stationary model identified by AIC sum to one, there is evidence of overdifferencing, and thus the trend stationary model alternative is selected. The selected forecast model for Rubber is ARIMA(1,0,0) with trend.

---

[12] The normalised drift coefficient estimate for the difference stationary model selected by SBC is also -0.105.

[13] It can generally be said that a unit root null hypothesis is preferable to a stationarity hypothesis in the particular case of forecast model selection, since the cost of misspecifying the order of integration of the data generating process tends to be lower for differenced models. (*Cf.* the simulation results quoted above. The standard approach of falling back on differenced models has also been a motivating factor for Diebold and Kilian's study (Diebold and Kilian (op. cit.)).

Sugar: This series is identified as I(0) by the ADF test and in this case would have a significant trend term. The simulation results on the impact of serial correlation on critical values in finite samples are not directly applicable to the ARIMA(1,0,1) model selected by SBC, but the low value of the AR(1) coefficient estimate (of 0.425) in conjunction with a t-statistic of -4.108 make a significant trend coefficient for the undifferenced model appear plausible. With the point estimates of the trend and drift components taking normalised values of -0.034 and -0.038 respectively, the model in levels could be expected to yield superior forecasts if a trend is included while for the I(1) model this may not be true. If the series were identified as difference stationary in the first place, the above simulation results suggest that better forecast results should be expected from a model without drift, even though the standard t-test identifies the drift term as significant when the forecast model is selected by AIC. (Naturally, these conclusions should be treated with some caution since the model parameterisations here are different.)

Given the estimated value of the AR(1) coefficient and the magnitude of the point estimate for the trend, it would also seem likely however, that in this case pre-testing can be expected to improve forecasts on average, even if there were no strong support in favour of a trend or drift coefficient. It is also the case that the moving average coefficient estimates for the model selected by AIC sum to one, thus providing further evidence of overdifferencing. The selected model therefore is an I(0) model with trend.

Timber: The price series for timber is identified as I(0) by the augmented Dickey-Fuller test. The moving average coefficient estimates of the difference stationary model selected by AIC sum to one, indicating overdifferencing. Under these circumstances and considering the normalised trend and drift coefficient estimates of 0.076 and 0.080[14] respectively it seems plausible that the model in levels would yield superior forecasts when a trend is included. If the data generating process were assumed to be difference stationary and the model were to be selected by AIC, there would now be no clear indication as to whether a drift should be included. The selected difference stationary model would be ARIMA(0,1,5) with a significant drift term. The simulation results for an ARIMA(0,1,1) model with a drift coefficient of comparable magnitude give no clear indication as to the whether including a drift term would improve forecast performance over the length of the forecast horizon considered. In addition, one should be aware that these simulation results are not directly comparable. The point estimate for the autoregressive coefficient is 0.680. In view of this value and the support for the trend and drift coefficient estimates, pre-testing can be expected to lead to improvements in forecast performance on average. Thus the optimal choice of forecast model is again a trend stationary specification. In this case the ARIMA (1,0,0) model with trend estimated in chapter 3 is chosen.

In addition to these commodities, where the presence of a trend component has been established with a high degree of confidence there are a number of price series where one can be confident about the absence of a significant trend term. In the following, model selection will be discussed for the commodity price series

[14] The value quoted is for the drift coefficient in the difference stationary model selected by AIC which appeared significant. The corresponding value for the model selected by SBC is 0.050.

were the absence of a trend or drift component has been established by the testing procedures in Chapter 4.

Coffee: This series -like most of the trendless examples- has been identified as difference stationary by the ADF test. In the difference stationary case it would be clear that forecasts from models without drift (*i.e.* a pure random walk, or alternatively an ARIMA(1,1,1) model in this case) should yield better results even if the point estimates for the drift and trend coefficient (0.009 and 0.018 respectively if normalised) were significant. It is also likely that model selection through pre-testing can improve forecast performance in this case, bearing in mind that the selected stationary model is ARIMA(1,0,0) with the autoregressive coefficient close to 0.8 in models with and without trend. In any case, the unit root and significance testing procedures would lead one to select a difference stationary (ARIMA(1,1,1)) model without drift[15].

Cocoa: Again the series identified by unit root and trend coefficient testing is a difference stationary model without drift. (ARIMA (2,1,0)). The insignificant normalised point estimate for the drift coefficient on Cocoa is -0.038 with the (also insignificant) normalised trend coefficient estimate equal to -0.012 (the t-ratios are -0.556 and -0.457 for trend and drift respectively). In so far as the simulation evidence on the inclusion of trend and drift coefficients can inform the decision on

---

[15]Considering an ARIMA(1,1,1) model instead of the random walk identified by SBC, the autoregressive coefficient estimate is 0.931 while the moving average coefficient is on the invertibility boundary and remains close to one, though not quite on the boundary, if the drift term is excluded. This would seem to be indicative of overdifferencing, a conclusion also supported by the Leybourne McCabe test. However, the ARIMA(0,1,2) model selected by AIC fails to indicate overdifferencing. The overall situation could then be characterised as one of uncertainty regarding the order of integration of the series. It has been argued above though, that in such a case one would opt for a differenced model, since this would lead to the smaller increase in forecast errors if the wrong forecast model were used.

model selection in this case, the difference stationary process again should probably be modelled without drift even if the drift coefficient estimate were significant. Since the ADF pre-test identifies the price series for Cocoa as difference stationary the question whether a model selected by pre-testing should be preferred over a differenced model does not arise in this case. The series is therefore modelled as ARIMA(2,1,0) without drift.

Tea: The price series is identified as difference stationary by the Augmented Dickey Fuller test. The normalised point estimates for the trend and drift coefficients are -0.046 and -0.058 respectively. The point estimates for both the trend and drift coefficient are identified as insignificant by a standard t-test with a nominal 5% rejection region, although the t-ratio on the estimate for the trend coefficient takes a value of -1.879 and thus is not as far from the asymptotic critical value of -1.96 as the drift coefficient with a t-ratio of -0.573. Given the estimated AR(1) coefficient of $\hat{\phi} = 0.883$ in the ARIMA(1,0,0) model and the point estimate for the trend coefficient, improvements in the performance of forecast models selected by pre-testing can be expected. It was shown in Chapter 4 that in the case of a random walk model selected by SBC in the difference stationary case, an ARIMA(1,1,1) parameterisation should be considered as an alternative. In the case of Tea, this does not yield a significant drift coefficient estimate and does not indicate overdifferencing if the model is re-estimated without drift. (The estimated coefficient on the moving average term is on the invertibility boundary though, if the drift term is included.) The price series for Tea is therefore best modelled as a pure random walk with an ARIMA(1,1,1) model without drift considered as an alternative.

Bananas: The normalised point estimates for the trend and drift coefficients for Banana prices are -0.012 and 0.004. The t-ratios on the coefficient estimates (-0.346 and 0.043) do not allow one to reject the null hypothesis of a zero trend coefficient at any of the conventionally employed critical values. In view of the Dickey Fuller test results obtained, as well as the estimated AR(1) coefficient value of 0.926 in the ARIMA(1,0,0) model selected by SBC, Banana prices seem to be best modelled as a pure random walk. Again, an ARIMA(1,1,1) model will be considered as an alternative, although there is no evidence of overdifferencing in this case.

Jute: For Jute, the normalised trend and drift coefficient estimates take values of -0.031 and -0.037 respectively. Both are insignificant following the standard t-test and, naturally would also appear insignificant when t-ratios are adjusted for serial correlation following Sun and Pantula (1999). In terms of forecast performance, the incorporation of a significant trend coefficient would be expected to improve forecast performance while for a drift coefficient of the same magnitude this is less likely to be the case. Here again, the estimated ARIMA(1,0,0) model with $\hat{\phi} = 0.848$ is close to the models used in the simulation experiments while the difference stationary alternative is not. It does however seem generally appropriate to be reluctant about including small drift coefficients into difference stationary forecast models. Given the lack of evidence in favour of a significant trend or drift term, this will be omitted from the forecast model.

The ADF test and the Leybourne McCabe test identify the series for Jute as difference stationary so that the question of selecting between differenced and pre-test models does not arise. Thus, the selected forecast model for Jute prices is an ARIMA (0,1,2) model without drift.

Tobacco: The ADF test clearly fails to reject the null hypothesis of a unit root for Tobacco. The normalised estimates for the trend and drift coefficient (0.033 and 0.019) are even lower in absolute value than in the case of Jute and both are clearly insignificant given t-ratios of 0.574 and 0.189 for the trend and drift coefficient estimates respectively. Considering further that the estimate of the autoregressive coefficient in the ARIMA(1,0,0) model specification takes a value of 0.953 it is obvious that the series is best modelled in first differences. This conclusion is supported by unit root test results as well as by the simulation evidence presented. The forecast model selected for Tobacco prices is therefore a pure random walk with an ARIMA(1,1,1) model considered as an alternative.

Copper: Again, the null hypothesis of a unit root could not be rejected by the ADF-test. The normalised trend and drift coefficient estimates take values of -0.023 and -0.050 respectively, with t-ratios of -1.097 and -0.490, while the estimated autoregressive coefficient in the ARIMA(1,0,0) model is $\hat{\phi} = 0.856$. Clearly, given the test and simulation results presented here, the price series is best modelled as difference stationary and without a drift component. In this case the selected difference stationary model is a pure random walk, while considering an ARIMA(1,1,1) model as an alternative.

Tin: The null hypothesis of a unit root can not be rejected for the price series of Tin. The normalised trend and drift coefficients for Tin are 0.005 and -0.017 respectively. Both are clearly identified as insignificant by their low t-ratios, and even have opposite signs in the point estimates. The autoregressive coefficient estimate for the ARIMA(1,0,0) model is $\hat{\phi} = 0.886$. The simulation results on pre-testing suggest that pre-testing for unit roots should improve forecasts on average, given the clear inference on the trend coefficient. The forecast model selected in this case is therefore once more the difference stationary model without drift, *i.e.* a pure random walk, and as before an ARIMA (1,1,1) model will be considered as an alternative.

Silver: The price series for silver is identified as difference stationary by the ADF test. The normalised trend and drift coefficients are estimated to take values of 0.001 and -0.018 with t-ratios of 0.050 and -0.229 respectively. Given the t-test on the trend coefficient estimates and the unit root test results, the obvious choice of forecast model is again the difference stationary alternative without drift, *i.e.* in this case an ARIMA(2,1,0) model.

Zinc: The ADF test clearly rejects the unit root null hypothesis here, while the Leybourne McCabe test fails to reject the null hypothesis of stationarity. As above in the case of Aluminium, there is strong evidence to suggest that the difference stationary model chosen by SBC is overdifferenced. The trend stationary model chosen by SBC is ARIMA(1,0,1):

[5.2.6] $$p_t = -0.009 + 0.001t + u_t, \; u_t - 0.427u_{t-1} = \varepsilon_t + 0.381\varepsilon_{t-1}$$

Differencing and defining $v_t = (1-L)u_t$ as above then yields:

[5.2.7] $$\Delta p_t = 0.001 + v_t,\ v_t - 0.427v_{t-1} = \varepsilon_t - 0.619\varepsilon_{t-1} - 0.381\varepsilon_{t-2}$$

By comparison, the estimated ARIMA(1,1,2) model identified by SBC is:

[5.2.8] $$\Delta p_t = 0.000 + v_t,\ v_t - 0.478v_{t-1} = \varepsilon_t - 0.639\varepsilon_{t-1} - 0.361\varepsilon_{t-2}$$

Again, the values are reasonably similar (the point estimate for the drift coefficient just rounds to 0.001 in 5.2.7). The estimates for the moving average coefficients in the ARIMA(1,1,2) model again sum to one.

The normalised point estimates for the trend and drift coefficient are 0.003 and 0.002 -both are clearly shown to be insignificant by the very low value of their t-ratios (0.369 and 0.184). The autoregressive coefficient value in the ARIMA(1,0,1) model for Zinc is estimated to be $\hat{\phi} = 0.427$. The simulation evidence is not directly applicable, but in view of the unit root and stationarity test results the ARIMA(1,0,1) model is chosen as a forecast model and re-estimated without trend.

In addition to the price series covered so far, there are a number of further series where there remains considerable uncertainty about the significance of the trend or drift coefficient. This is most often the case where the result of the standard t-test depends on *a priori* assumptions on the order of integration and where furthermore this question is not resolved by the supplementary testing procedures described in Chapter 4.

Rice: Among the various commodities with some remaining uncertainty over the significance of the trend coefficient, this is perhaps the case closest to the group for which the presence of a trend was inferred with a high degree of confidence.

In the case of Rice, the trend and drift coefficients (with normalised point estimates of -0.069 and -0.072 respectively) are identified as significant by their respective t-ratios, even if the difference stationary model is selected by SBC. One could therefore argue that the original motive for turning to complementary tests should be confined to considering the impact of serial correlation if the series is modelled in levels. The simulation results and bias corrections proposed by Sun and Pantula are not directly applicable here, because of the ARIMA(1,0,1) parameterisation for the model in levels. (The t-ratio on the estimated trend coefficient for Rice prices (-5.223) is, however, large enough to require highly persistent serial correlation for a spurious rejection of the null hypothesis of a zero trend coefficient to be a likely problem.) Vogelsang's test by contrast provides no support for the presence of a trend, although this also tends to be the case for a range of series which appear to have a trend when judged by any of the other criteria available and may simply illustrate the low power of the test in the case of integrated series.

Regarding the order of integration of the series the evidence from pre-testing indicates that a difference stationary model is most appropriate since the ADF-test fails to reject the null hypothesis of a unit root and the Leybourne MacCabe test rejects the stationarity null hypothesis. There is no evidence of overdifferencing from either the difference stationary model selected by either SBC or AIC. Opting for a differenced model in the case of persistent uncertainty about the order of integration of the data generating process, the selected model is an ARIMA(1,1,2) model with drift.

Wheat: The price series for Wheat is identified as I(1) by the ADF test. (The Leybourne McCabe test fails to reject the stationarity hypothesis, although in such a case of uncertain inference about the order of integration, the conservative choice -supported by the above simulation experiments- is in favour of a difference stationary forecast model.)

The normalised point estimates for the trend and drift coefficient are -0.069 and -0.062 with t-ratios of -6.866 and -1.154 respectively. Of these, the t-test suggests that the coefficient estimate is significant for the model in levels but not for the model in first differences (as usual, requiring rejection of the null hypothesis with 95% confidence in a two tailed test). When the difference stationary model is selected by AIC, the evidence against the drift coefficient is weakened substantially with the t-ratio on the drift coefficient falling to -1.923. Of the Vogelsang test statistics, only one rejects the null hypothesis of a zero trend coefficient, while the Sun and Pantula corrections are not applicable here. At the point estimates for the trend and drift coefficient, pre-testing would be expected to yield improved forecast results for any but near integrated first order autoregressive series. The magnitude of the point estimates for the trend and drift coefficients make it seem unlikely that omission of the trend coefficient would lead to improved forecast results, while this is less clear for difference stationary models. (Given the different model parameterisation in this case one should, of course be careful in generalising the implications of the simulation results.) Since the t-ratio on the drift coefficient rises substantially in the model selected by AIC *vis a vis* the one in the model selected by minimum SBC and in view of the fact that the t-ratio is close to the 5% critical value, it seems appropriate to select a forecast model with drift. The selected forecast model in this case is ARIMA(0,1,4) with drift. (The alternative forecast model without drift, selected by SBC is presented in Appendix V.ii.)

Maize: The point estimates for the normalised trend and drift coefficients for Maize are -0.049 and -0.048 respectively. If one were to apply the above simulation results on the performance of forecast models with and without trend or drift coefficient correctly including a significant trend would be likely to improve the forecast performance of a trend stationary model. The trend stationary model in this case is ARIMA(1,0,0) with $\hat{\phi} = 0.720$, the difference stationary ARIMA(1,1,2) model is much more different from the model parameterisation used in the simulation experiment[16]. The simulation evidence on the difference stationary model does suggest however, that models without drift have smaller mean squared forecast errors on average. In any case, the t-ratios would identify the trend coefficient as significant for the I(0) model only.

Neither Vogelsang's test nor the Sun and Pantula results lend further support for the incorporation of a trend coefficient. The unit root test supports a difference stationary model, as does the Leybourne McCabe test, there is no evidence of overdifferencing and a comparison of the forecast errors, on the basis of the simulation results quoted earlier, shows that wrongly fitting a difference stationary

[16] The point estimate of the autoregressive coefficient reported here has not been used in the simulation experiments for stationary models either. However, the model parameterisation corresponds to the one used in the simulation experiments and in so far as inference on model performance for different magnitudes of the autoregressive coefficient is possible, it appears that a lower trend coefficient estimate is more likely to yield superior forecast results for the lower autoregressive coefficient. While one should exercise some caution in generalising the simulation results, the inclusion of a trend term would not appear implausible in this case.

model without drift tends to lead to a smaller increase in forecast errors than wrongly fitting a trend stationary model. In contrast to Wheat, alternative model selection by AIC does not yield stronger evidence in favour of a drift coefficient. The most appropriate forecast model therefore seems to be an ARIMA(0,1,2) model, re-estimated excluding the drift coefficient.

Beef: Here the normalised point estimates for the trend and drift coefficient are 0.067 and 0.039 respectively. As in a large number of other cases, the trend coefficient would appear significant at the standard critical value for the t-test when the model in levels is considered, but not for a random walk plus drift. The Augmented Dickey Fuller test identifies the price series as I(1). If critical values are modelled in simulation experiments -as in Sun and Pantula (1999) and above in Chapter 4- based on the assumption that the underlying data generating process is difference stationary, then the t-ratio would no longer suggest that the trend coefficient is significant. This conclusion still holds if the series is modelled as a trend stationary first order autoregressive process with $\hat{\phi} = 0.905$, since in this case the adjusted 5% critical value as in table 4.1.3. would be ±3.29 while the actual t-ratio on the trend coefficient is $t_{\hat{\beta}} = 2.295$.

Given the normalised point estimate on the trend coefficient, including a significant trend coefficient should lead to improved forecast performance in a trend stationary model. (The ARIMA (1,0,0) model with an estimated AR(1) coefficient value of 0.905 allows for a comparison with the relevant simulation results, in contrast to the difference stationary case.) This may not be the case for the difference stationary alternative, where better forecasts could be obtained without a drift term even if it was significant. Since the unit root pre-test would indicate the selection of a difference stationary model in the present case and since there is evidence in support of a trendless model in either case, the most appropriate forecast model for relative Beef prices seems to be a pure random walk, or an alternative driftless ARIMA(1,1,1) model.

Lamb: The unit root hypothesis is rejected by the Augmented Dickey Fuller test for this series. The normalised point estimates obtained for the trend and drift coefficient respectively are 0.094 and 0.072. The t-ratios are 5.102 and 0.715 respectively. Standard critical values for the t-ratio would suggest a significant trend and insignificant drift term. It is moreover worth noting that the point estimate for the trend coefficient is larger than the corresponding estimate for the drift term. One could therefore expect an improvement in forecast performance when correctly including the trend term to be more likely than in the case of the drift term, although the model parameterisations for Lamb do not allow for a direct comparison with the simulation evidence on forecast performance. The evaluation of the significance of the trend component is therefore crucial for the quality of the forecasts obtained. The issue is complicated by the fact that the Sun and Pantula corrections are not directly applicable, while of the Vogelsang test statistics two indicate the presence of a trend while the other two do not. (Given the low power of the Vogelsang test in this case, this evidence is however worth considering.) The fact that the ADF test as well as the Leybourne McCabe test provide evidence of a stationary model should justify the selection of a stationary forecast model[17].

[17] If an ARIMA(1,1,1) model is considered as an alternative to the random walk selected by SBC, the estimated coefficient on the moving average term is on the invertibility boundary, though this is

In view of the large t-statistic on the trend coefficient, the inclusion of a trend component also seems appropriate. In spite of the remaining factors of uncertainty, an ARIMA(5,0,0) forecast model with trend is therefore selected.

Palm Oil: The normalised point estimates for the trend and drift coefficients are -0.049 and -0.033 respectively. The series is identified as difference stationary by the ADF test while the Leybourne McCabe test fails to reject the stationarity hypothesis. A t-test at the conventional critical values would again indicate the presence of a significant trend coefficient for the model in levels while the drift coefficient with a t-ratio of -0.429 would not be considered statistically significant. The t-statistic on the trend coefficient is -4.332 and although the Sun and Pantula results are not directly comparable, the low first order autoregressive coefficient value of 0.591 suggests that the impact of serial correlation is unlikely to raise doubts about the adequacy of standard critical values if the series were to be modelled in levels. Among the Vogelsang test statistics, there is no support for the presence of a significant trend.

The simulation evidence on the comparative forecast performance of models with and without trend is not directly applicable. If one were to judge the issue on the basis of the magnitude of the point estimates for trend and drift coefficients, however, it would appear that improved forecasts could result from the inclusion of a trend term while for the difference stationary alternative models without drift would be preferred. (One should remember though that the model parameterisation underlying the simulation results is different from the one in the estimated models

---

not the case for the difference stationary alternative identified by AIC. This provides some additional support for a trend stationary model.

for Palm Oil.) Hence, considering all the evidence available, the selection of the most appropriate forecast model should depend either on pre-testing or on the lowest cost of a misspecified model. Given the contradictory evidence from unit root and stationarity tests and the ensuing uncertainty in inferring the appropriate order of integration, a difference stationary model without drift is the preferred option here, given the lower expected cost of misspecification as discussed above. The forecast model used therefore is an ARIMA(2,1,0) model as identified by SBC and re-estimated without drift.

Cotton: Here the ADF test fails to reject the unit root null hypothesis. The normalised point estimates for the trend and drift coefficients respectively are -0.061 and -0.053. The t-ratio on the trend coefficient would suggest that the trend is significant at standard critical values, although the autoregressive coefficient in the ARIMA(1,0,0) model takes a value of $\hat{\phi} = 0.833$ so that the impact of serial correlation should be sufficient to be cautious about the use of standard critical values even if the series were to be modelled in levels. However, a priori conclusions on the order of integration are important here. The t-ratio obtained, -3.328 would be significant at standard asymptotic critical values as well as if serial correlation is taken into account for the estimated value of the first order autoregressive coefficient: the adjusted critical values shown in table 4.1.3 are -2.65 for $\phi = 0.8$ and -2.87 for $\phi = 0.85$. Conclusions on significance would be reversed however, if the critical values were to be adjusted for a difference stationary process. Thus the choice of model selection once again is crucially dependent on pre-testing.

The assumption of an integrated series seems appropriate for forecast model selection however. The simulation experiments on expected improvements from pre-testing suggest that at the given point estimates for the trend and drift coefficients performance improvements from pre-testing may be possible (although the parameterisation of the difference stationary alternative differs from the one used in the simulation experiments, so some caution is in order). A driftless I(1) model would obviously also be favoured if a choice had to be made in the presence of persistent uncertainty surrounding the use of pre-tests. Given that the ADF test as well as the Leybourne McCabe test support a differenced model, this does not seem to be a problem here. In so far as the results of the simulation experiments on forecast performance with and without trend or drift coefficient can be generalised, the inclusion of a trend coefficient is again more likely to be of advantage in the trend stationary than in the difference stationary model and the t-ratio on the drift coefficient estimate remains well below the standard critical value of ±1.96 for all the difference stationary alternatives considered. The forecast model chosen for Cotton prices is therefore the ARIMA(2,1,2) model identified by SBC and re-estimated without drift.

Wool: The trend and drift coefficient estimates for this price series have normalised values of -0.081 and -0.078 respectively. On the basis of standard t-tests, the coefficient estimate for the trend stationary model (with a t-ratio of -4.652) would appear significant while the drift coefficient estimate with a t-ratio of -1.850 would be seen as insignificant at the conventional critical value of -1.96. This latter result should be interpreted with care since the P-value for the t-ratio on the drift coefficient for wool is 0.067, thus being reasonably close to 5%. The t-ratio on the trend coefficient in the trend stationary model would not appear significant if the critical values are adjusted for an underlying difference stationary process following Sun and Pantula (1999). It would however still appear to be significant when the estimate of the autoregressive coefficient were to be taken at face value with $\hat{\phi} = 0.824$ and an implicit adjusted critical value between -2.65 and -2.87 (from table 4.1.3.).

The series is identified as I(1) by the ADF test. The magnitude of the trend coefficient estimates suggests that the inclusion of a significant trend coefficient would improve the forecast performance of a model in levels. (Since the model selected on the basis of the SBC is an ARIMA(1,0,0) model with an AR(1) coefficient estimate of 0.824, the comparison of the trend stationary model with the simulation results is relatively unproblematic. For the case of a model in first differences, the situation is less clear, with respect to the simulation results on I(1) DGPs as well as because the model parameterisation differs from the one employed in the simulation experiments. In the simulation results reported above, difference stationary data generating processes with a moving average coefficient of $\theta = 0.2$ benefit from the inclusion of a significant drift term with a value of $\beta = 0.07$ and above, from a forecast horizon length of h=2 and beyond. This is not the case where $\theta = 0.1$, since here the value for the drift coefficient estimate required for superior forecasts is of at least $\beta = 0.09$, in which case improved forecast results will result at a forecast horizon length of h=9 and beyond.

The simulation results on pre-testing suggest that improvements could be expected from pre-testing for unit roots in this case -regardless on whether the inclusion of a trend or drift coefficient depends on stationarity assumptions. Regarding the assumed order of integration, it seems again appropriate to use a difference stationary forecast model. One remaining problem is of course the continuing uncertainty surrounding the inference on the inclusion of a drift term in the difference stationary model. There is also some uncertainty as to whether the inclusion of a significant drift term would improve expected forecast results given the point estimates obtained.

Once more then, it can be concluded that the most appropriate forecast model is a difference stationary model without drift: What conclusions can be drawn from the simulation evidence available suggests that difference stationary models are less likely to perform well when small drift coefficients are incorporated. It has therefore been decided to drop the drift term although some uncertainty remains. The selected forecast model is an ARIMA(0,1,2) model re-estimated without drift.

Lead: The normalised point estimates for the trend and drift coefficient for Lead are -0.033 and -0.042 respectively. Thus inclusion of a significant trend term should lead to better forecast results for a forecast model in levels but probably not for a forecast model in first differences. The estimated ARIMA(1,0,0) model with $\hat{\phi} = 0.795$ is close to the simulated model, but the difference stationary model (a random walk with drift) is not, so that the results on the performance of I(1) models with drift should be treated with some caution. The ADF test fails to reject the null hypothesis of a unit root, but the Leybourne McCabe test would identify the series as stationary. On the other hand, simulation results show that at the given point estimates for the trend and drift and autoregressive coefficients improvements in the average forecast performance of the selected model are to be expected if a pre-test is used.

Regarding the question of the significance of the trend coefficient, it should be noted that the trend coefficient appears significant only on the basis of a standard t-test when the series is modelled in levels. This conclusion no longer holds for the model in first differences or on the basis of the Sun and Pantula or Vogelsang tests. The trend coefficient estimate also no longer appears significant if critical values are adjusted for serial correlation at the estimated value of the autoregressive coefficient $\hat{\phi} = 0.795$. There is thus substantial evidence against the inclusion of a trend or drift coefficient in the forecast model for Lead. Given the lack of support for the inclusion of a trend coefficient and the remaining uncertainty surrounding the application of pre-tests in this case it appears appropriate to model the price series for Lead as difference stationary without drift. It is possible that the true generating process may be stationary in levels in spite of the pre-test results (given the outcome of the Leybourne-McCabe test).

Imposing an ARIMA(1,1,1) model reveals a moving average coefficient estimate on the invertibility boundary[18] without leading to a substantial drop in the standard error on the drift coefficient estimate. (This finding does not support the hypothesis that the drift coefficient estimate appears insignificant because the difference stationary model alternative is underparameterised.) The autoregressive coefficient

---

[18] However this is not the case for the difference stationary model selected by AIC.

estimate in the ARIMA(1,1,1) model is clearly significant, considering a t-ratio of 7.949 while a meaningful standard error on the moving average coefficient estimate can not be obtained since the point estimate is on the invertibility boundary.

In conclusion, a case for modelling the price series for Lead as a trendless stationary process could clearly be made. However, since the insignificance of the trend coefficient estimate has also been established, the absence of a trend coefficient can be taken as given when assessing the benefits of unit root pre-testing for the selection of the appropriate forecast model. It has been shown above that the average performance of a forecast model selected thus should be expected to improve in the present case (*i.e.* given the estimate of the autoregressive coefficient and the absence of a trend component regardless of the inferred order of integration.) In addition, it is also to be expected that the cost of misspecification is generally lower when a difference stationary alternative is used (*Cf.* footnote 13 above, where model selection for Rubber is discussed). The selected forecast model for Lead is therefore an ARIMA(1,1,1) model (considering a pure random walk as a possible alternative).

estimate in the ARIMA(1,1,1) model is clearly significant, considering a t-ratio of 7.949 while a meaningful standard error on the moving average coefficient estimate can not be obtained since the point estimate is on the invertibility boundary.

In conclusion, a case for modelling the price series for Lead as a trendless stationary process could clearly be made. However, since the insignificance of the trend coefficient estimate has also been established, the absence of a trend coefficient can be taken as given when assessing the benefits of unit root pre-testing for the selection of the appropriate forecast model. It has been shown above that the average performance of a forecast model selected thus should be expected to improve in the present case (*i.e.* given the estimate of the autoregressive coefficient and the absence of a trend component regardless of the inferred order of integration.) In addition, it is also to be expected that the cost of misspecification is generally lower when a difference stationary alternative is used (*Cf.* footnote 13 above, where model selection for Rubber is discussed). The selected forecast model for Lead is therefore an ARIMA(1,1,1) model (considering a pure random walk as a possible alternative).

5.3. Relative Price Forecasts

5.3.1. Price forecasts obtained

An overview of the forecast models selected is given in table 5.3.1 below. The forecast results obtained are listed below in tables 5.3.2. to 5.3.7 for a ten period forecast horizon and for point forecasts only. Ten period forecasts with 68% confidence intervals are listed in Appendix V.i, where confidence intervals are calculated on the basis of the estimated forecast model, abstracting from additional uncertainty regarding the accuracy of model specification. The graphs presented in the text plot forecasts and 68% confidence intervals over 20 years, to show the ten year projections in context.

Table 5.3.1. Forecast models used for individual commodity price series.

Commodity	Forecast model*	Trend /Drift	Commodity	Forecast model*	Trend /Drift
Aluminium	1,0,1	Y	Silver	2,1,0	N
Rubber	1,0,0	Y	Zinc	1,0,1	N
Sugar	1,0,1	Y	Rice	1,1,2	Y
Timber	1,0,0	Y	Wheat	0,1,4	Y
Coffee	1,1,1	N	Maize	0,1,2	N
Cocoa	2,1,0	N	Beef	0,1,0	N
Tea	0,1,0	N	Lamb	5,0,0	Y
Bananas	0,1,0	N	Palm Oil	2,1,0	N
Jute	0,1,2	N	Cotton	2,1,2	N
Tobacco	0,1,0	N	Wool	0,1,2	N
Copper	0,1,0	N	Lead	1,1,1	N
Tin	0,1,0	N			

ARIMA(p,d,q) specifications are given for forecast models.

In the cases where difference stationary forecast models without drift have been selected, Appendix V.iii. lists the re-estimated ARIMA models without drift. In those cases where the selected forecast model is a pure random walk, the model estimates presented are for an ARIMA(1,1,1) model without drift.

One point worth mentioning is that the forecast models selected for Wheat and Maize differ more than one might expect, given that the two commodities should be substitutes to some degree. Although both price series are identified as difference stationary, only the forecast model for Wheat contains a drift term. One should remember here that the case for a drift term was far from clear cut but that the time series characteristics of the price series for Wheat were such as to motivate the inclusion of a drift term. Similar considerations may arise in the case of Lamb and Beef.

Table 5.3.2. below shows forecasts up to ten periods ahead for the four commodities for which the presence of a trend or drift coefficient was regarded as highly likely.

Table 5.3.2. Point forecasts for some models with trend or drift

h	**Aluminium** ARIMA(1,0,1)	**Rubber** ARIMA(1,0,0)	**Sugar** ARIMA(1,0,1)	**Timber** ARIMA(1,0,0)
0	*-0.394*	*-0.931*	*-0.540*	*-0.282*
1	-0.508	-0.826	-0.290	-0.142
2	-0.529	-0.854	-0.301	-0.043
3	-0.550	-0.882	-0.311	0.028
4	-0.569	-0.910	-0.321	0.080
5	-0.589	-0.938	-0.332	0.119
6	-0.608	-0.966	-0.342	0.149
7	-0.627	-0.994	-0.353	0.173
8	-0.646	-1.022	-0.363	0.193
9	-0.665	-1.050	-0.373	0.210
10	-0.683	-1.078	-0.384	0.225

h: Forecast horizon, *h*=0 last observation of the original data set, All forecasts are for prices relative to MUV in natural logarithms.

The results reported in table 5.3.2. reflect the presence of a negative trend coefficient in the cases of Aluminium, Rubber and Sugar and of a positive trend in the case of Timber. The forecast values fall or rise accordingly. The corresponding point forecasts and 68% confidence intervals for a forecast period of 20 years are illustrated in Figures 5.3.1 to 5.3.4 below.

Figure 5.3.1

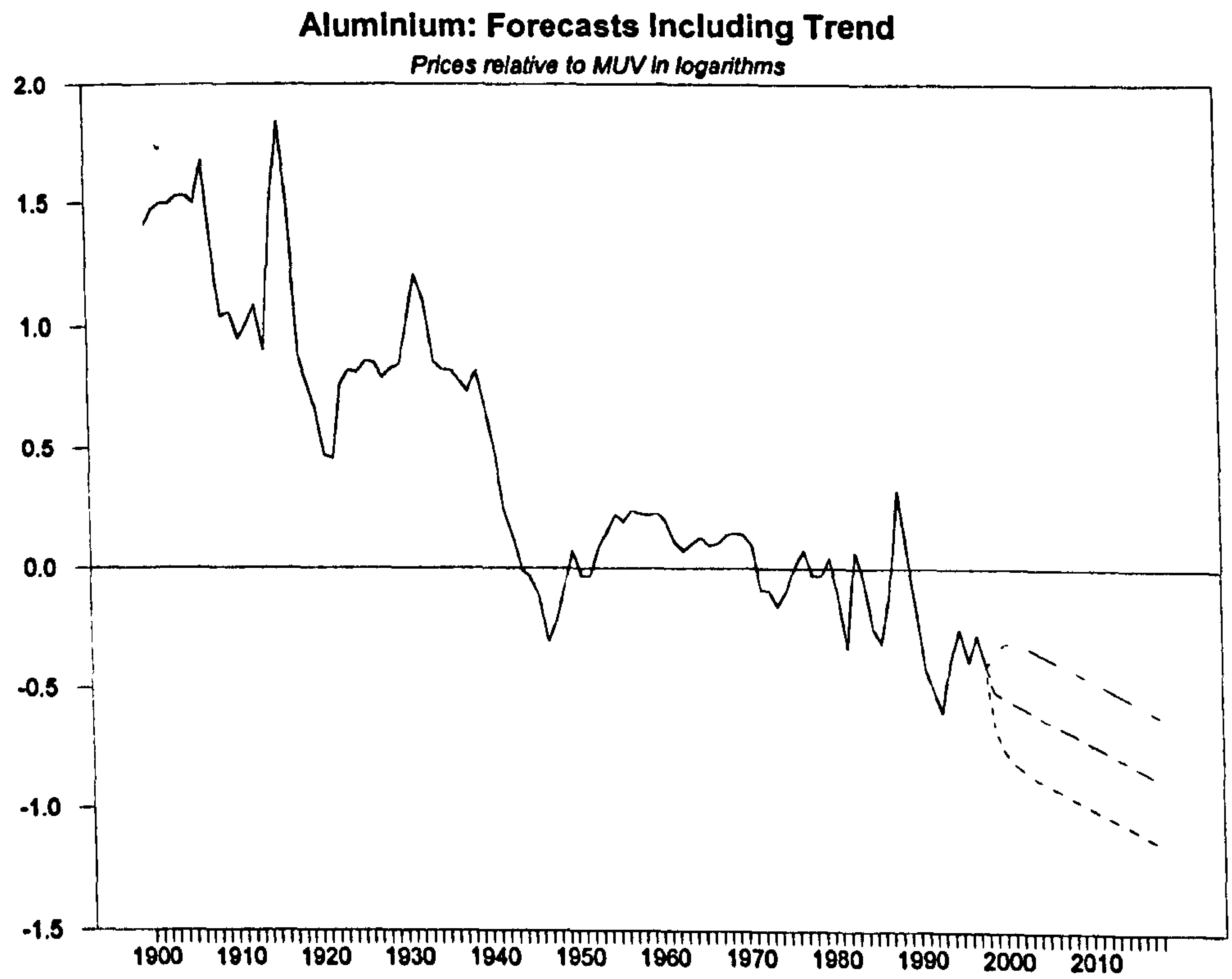


Figure 5.3.2

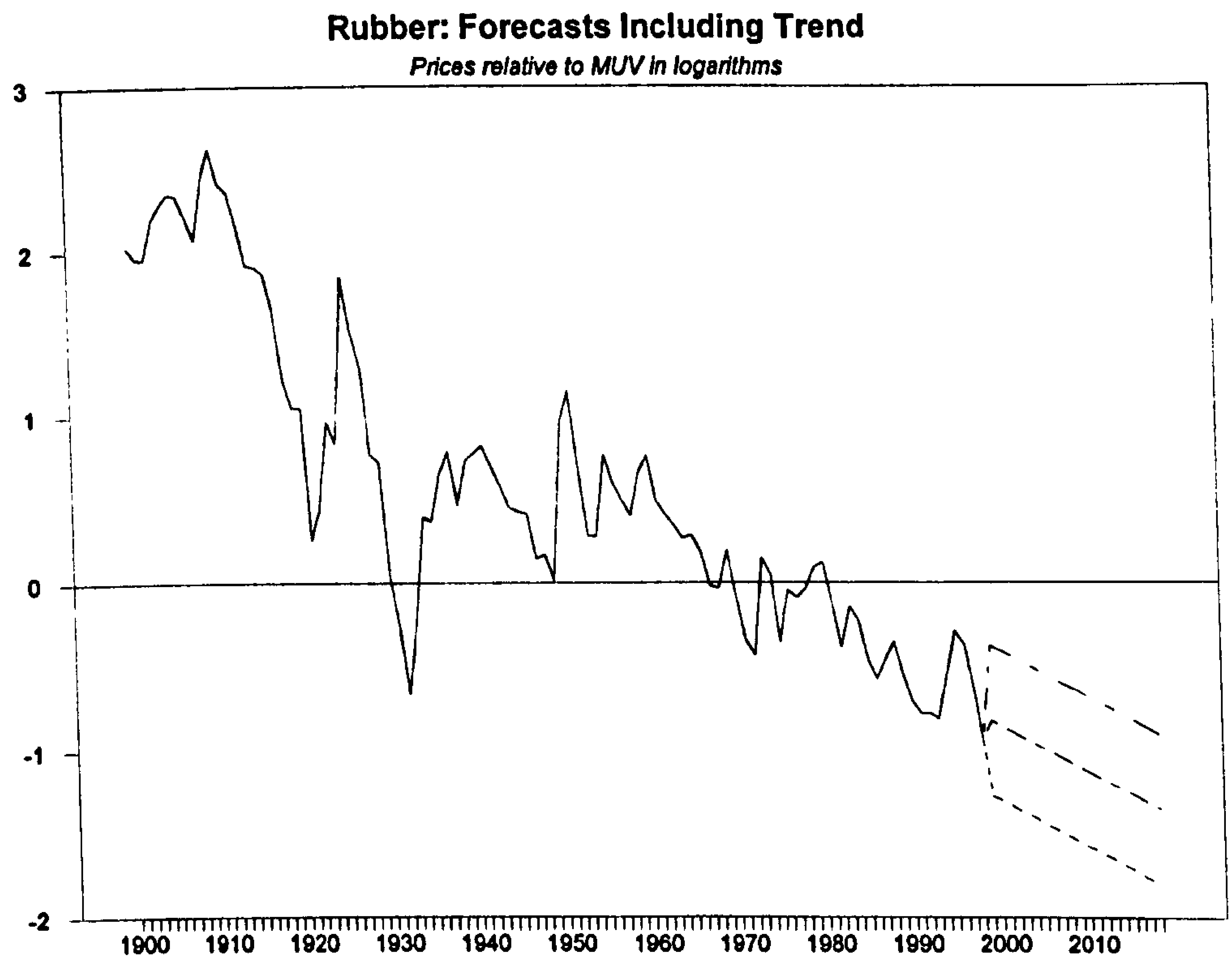


Figure 5.3.3:

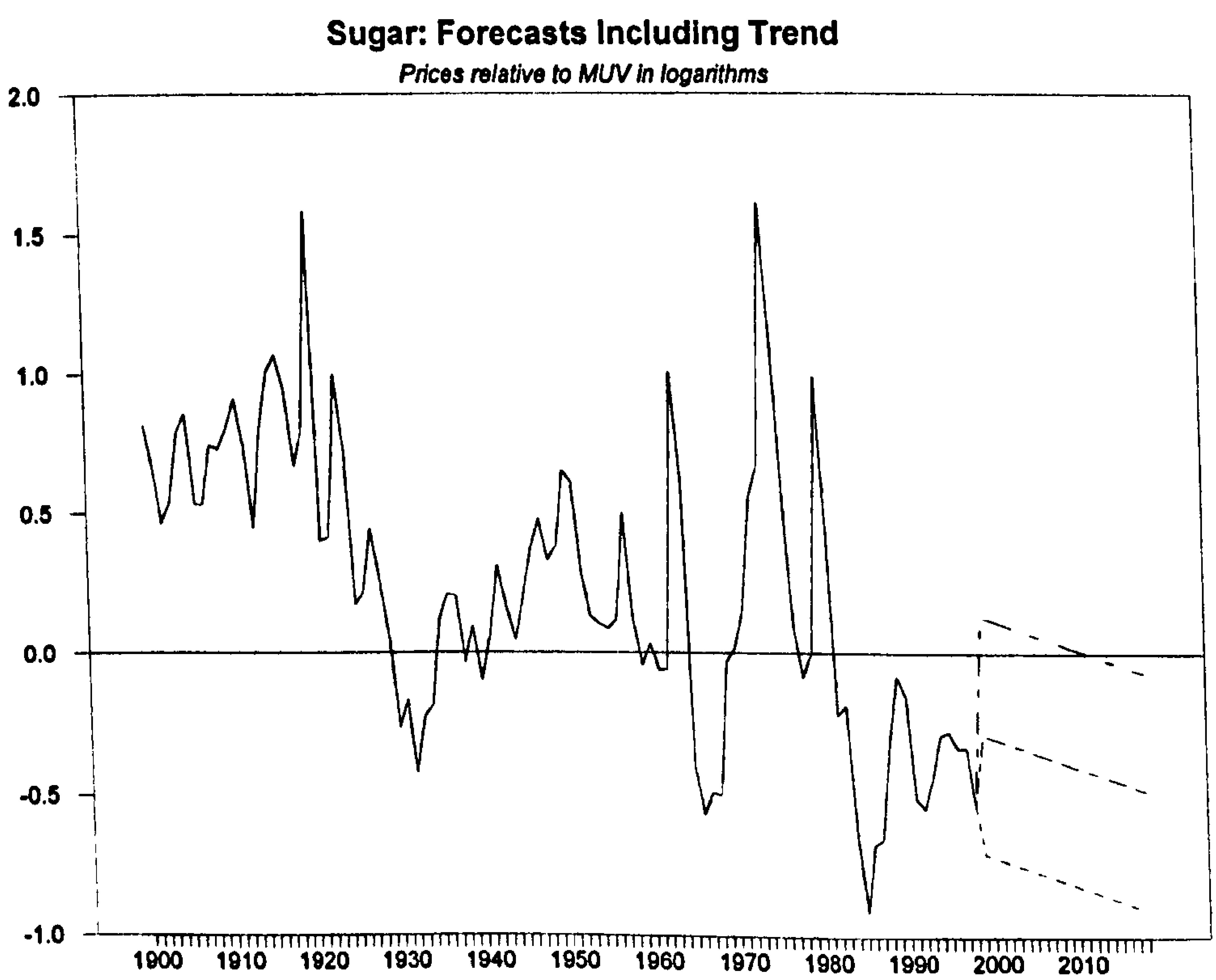
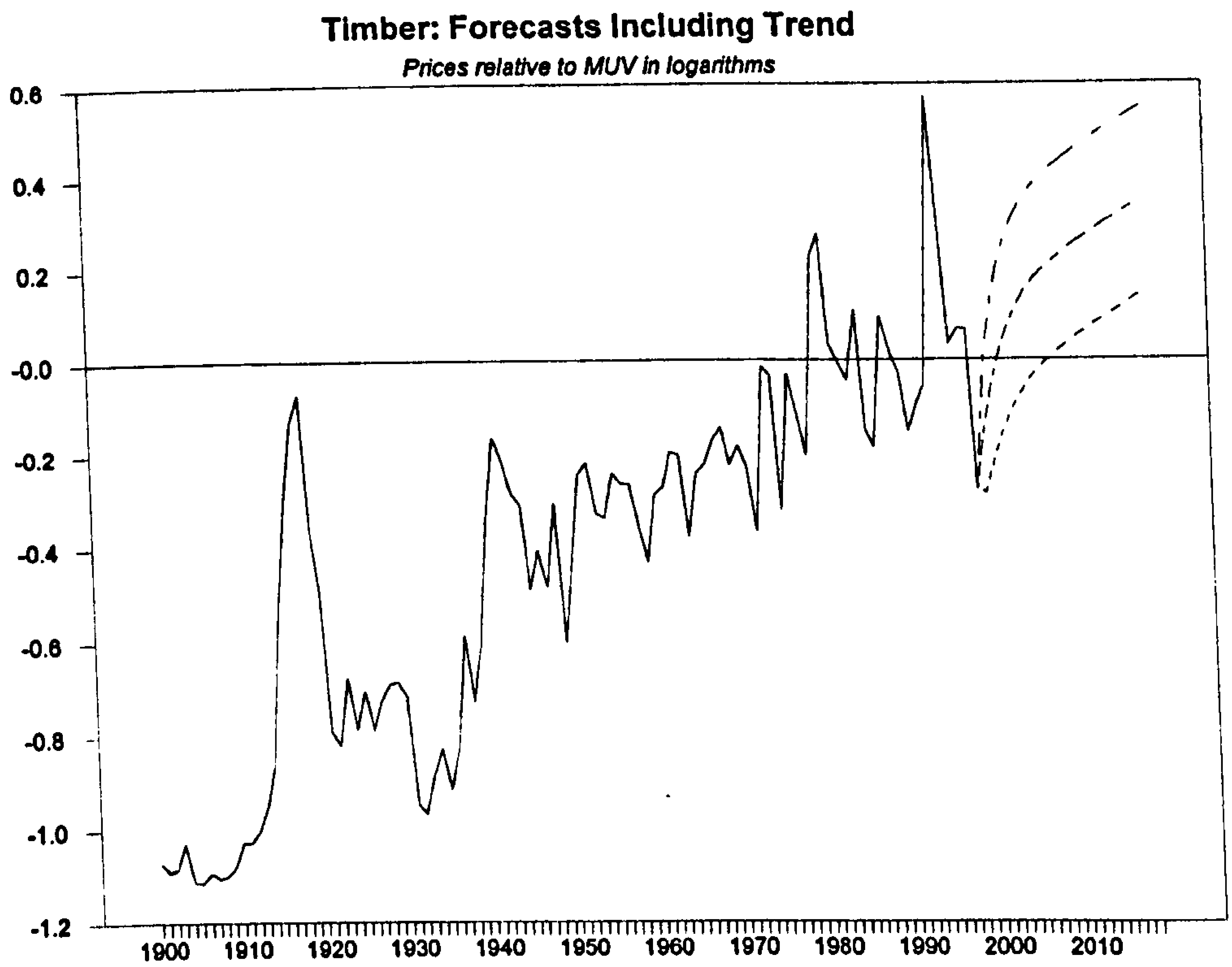


Figure 5.3.4:



For those commodities for which it was concluded that no trend or drift component is present the forecasts for those cases in which the forecast model selected is not a pure random walk or ARIMA(1,1,1) model are shown below in table 5.3.3.

Table 5.3.3. Point forecasts for models without trend or drift

h	**Cocoa** ARIMA(2,1,0)	**Jute** ARIMA(0,1,2)	**Silver** ARIMA(2,1,0)	**Zinc** ARIMA(1,0,1)
0	*-1.319*	*-0.942*	*-0.825*	*-0.117*
1	-1.363	-0.813	-0.817	-0.131
2	-1.389	-0.795	-0.866	-0.045
3	-1.378	-0.795	-0.871	-0.008
4	-1.369	-0.795	-0.856	0.008
5	-1.371	-0.795	-0.854	0.015
6	-1.374	-0.795	-0.858	0.018
7	-1.374	-0.795	-0.859	0.019
8	-1.373	-0.795	-0.858	0.020
9	-1.373	-0.795	-0.857	0.020
10	-1.373	-0.795	-0.858	0.020

h: Forecast horizon, *h*=0 last observation of the original data set, All forecasts are for prices relative to MUV in natural logarithms.

The trendless, mostly difference stationary forecast models in table 5.3.3. above yield forecasts that converge to a constant value after a short horizon. As one would expect, this convergence occurs more quickly for pure moving average processes than for models containing an autoregressive component. In the case of Zinc the forecasts converge on the unconditional mean of the series, while for the difference stationary models the constant forecast value obtained finally depends on the last observation in the original series and any inference from the ARMA components of the model[19]. The forecasts listed in Table 5.3.3 are illustrated over a period of 20 years and with 68% confidence intervals in figures 5.3.5 to 5.3.8 below:

[19] Where the 68% confidence intervals for the forecasts are shown in Appendix V.i. and in this Chapter it also becomes apparent that the confidence intervals for forecasts from stationary or trend stationary models are characterised by a bounded variance whereas the variance, and hence the confidence interval width, for integrated processes increases with the forecast horizon.

Figure 5.3.5:

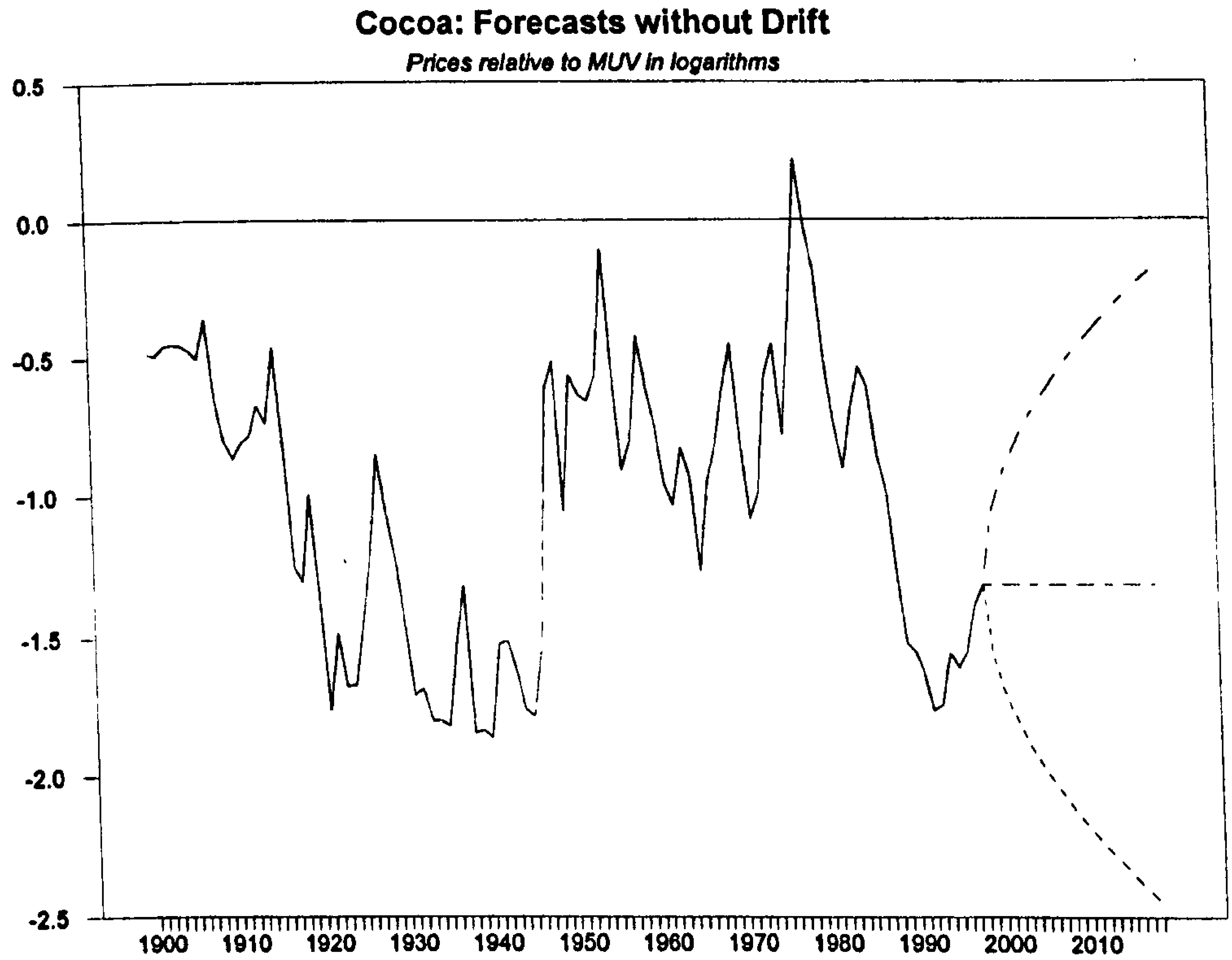


Figure 5.3.6:

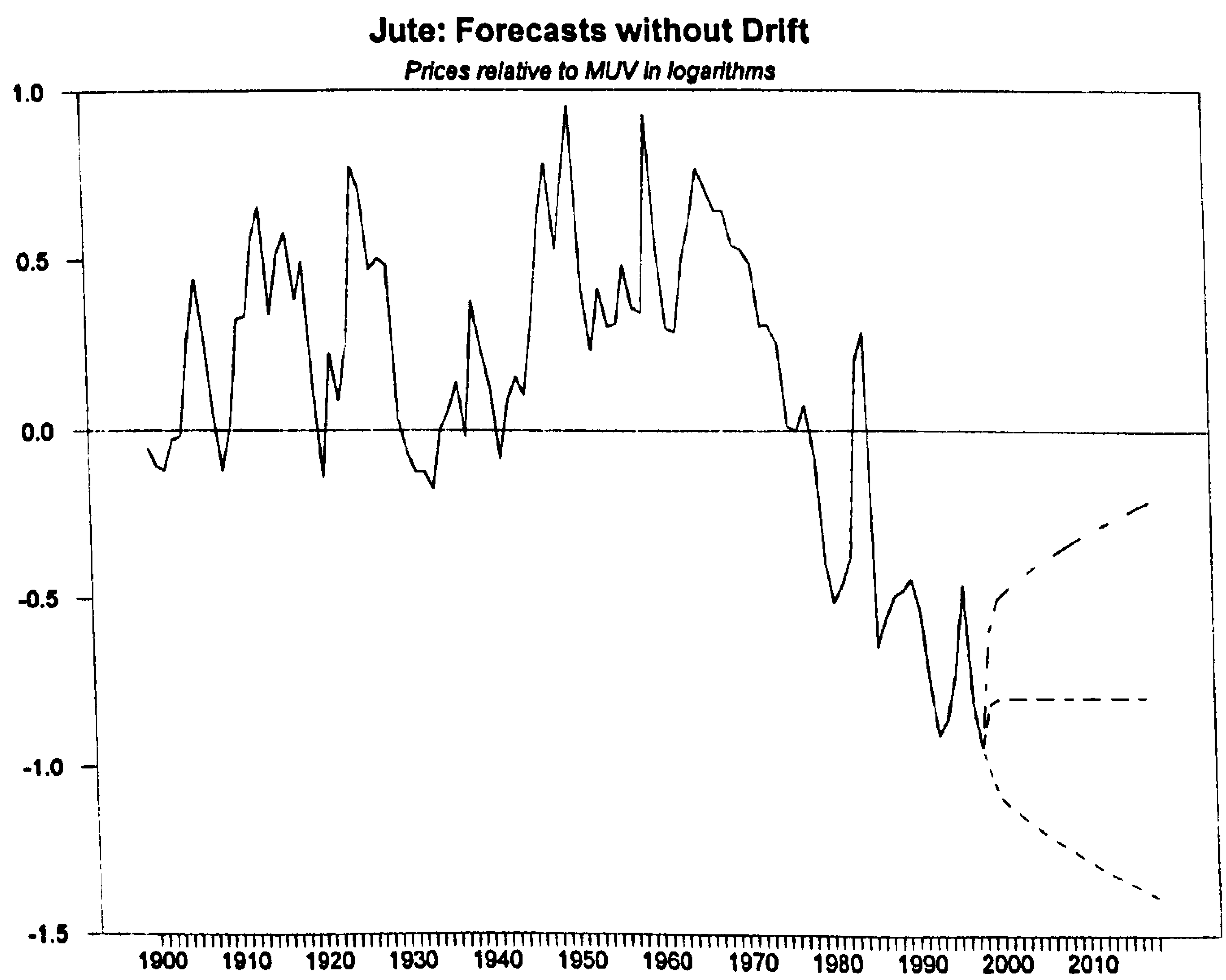


Figure 5.3.7:

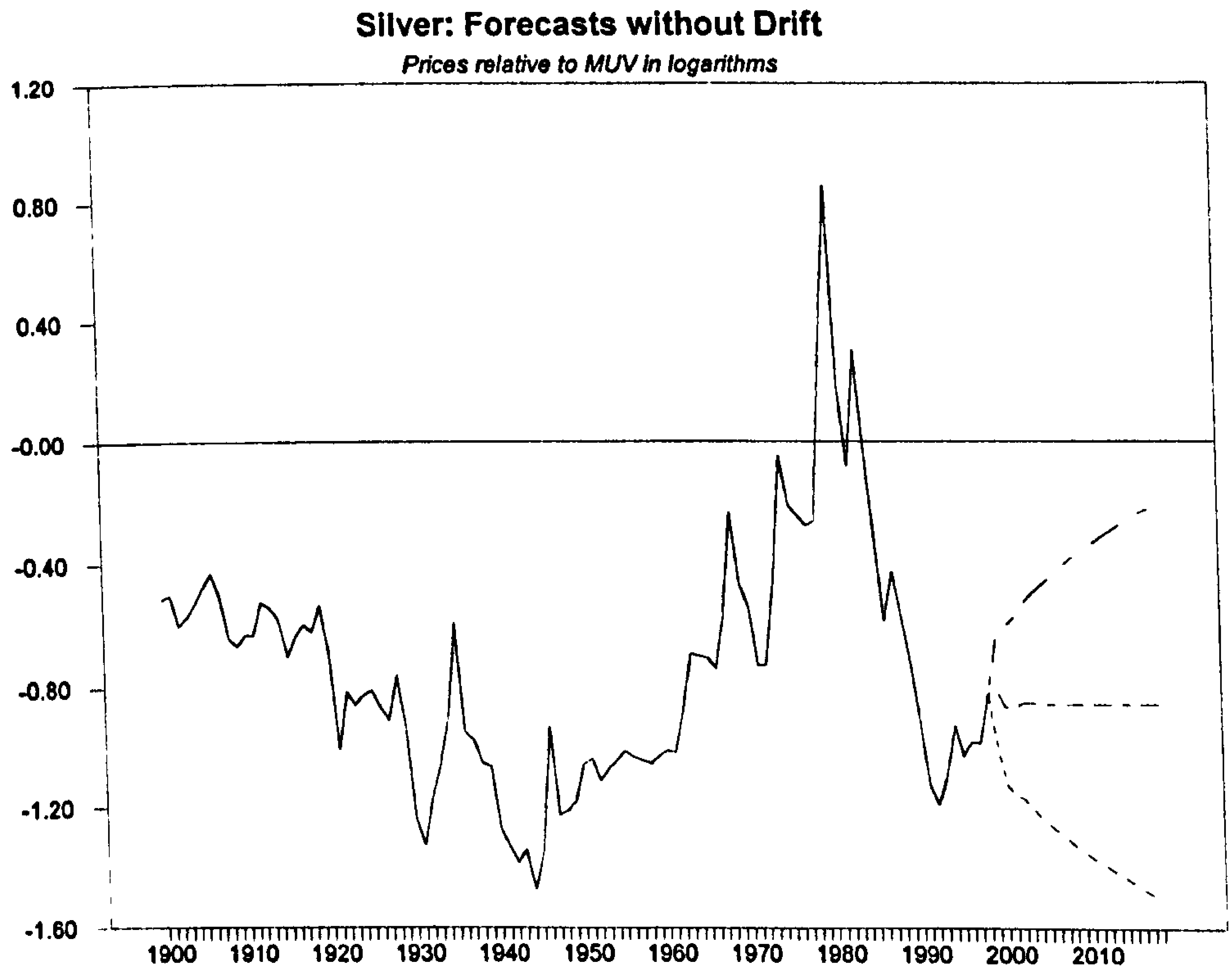
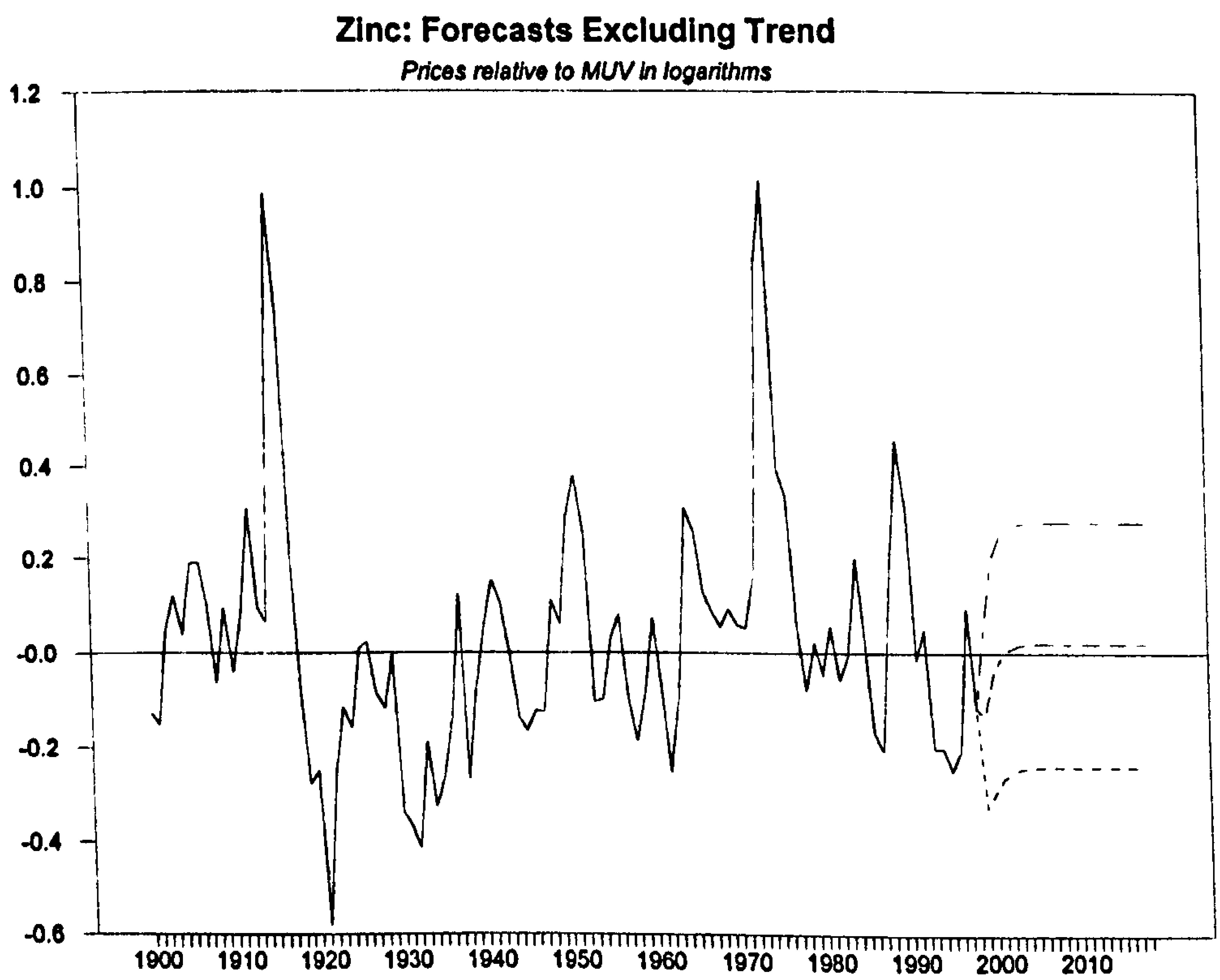


Figure 5.3.8:



Point forecasts from pure random walk models are simply the last observation in the original data series at any forecast horizon. Table 5.3.4. lists these values for all driftless models which would be modelled as a pure random walk, when selecting by SBC.

Table 5.3.4. Point forecasts from pure random walks

Commodity	Forecast	Commodity	Forecast
Coffee	-0.926	Tobacco	-0.898
Tea	-0.672	Copper	-0.522
Bananas	-0.077	Tin	-1.441

All forecasts in the table are from random walk models and are therefore identical for any forecast horizon. All forecasts are for prices relative to MUV in natural logarithms.

Forecasts from random walk models are illustrated in Figures 5.3.9-5.3.13 below. (The forecasts for Coffee will be illustrated below when ARIMA(1,1,1) models are presented.

Figure 5.3.9:

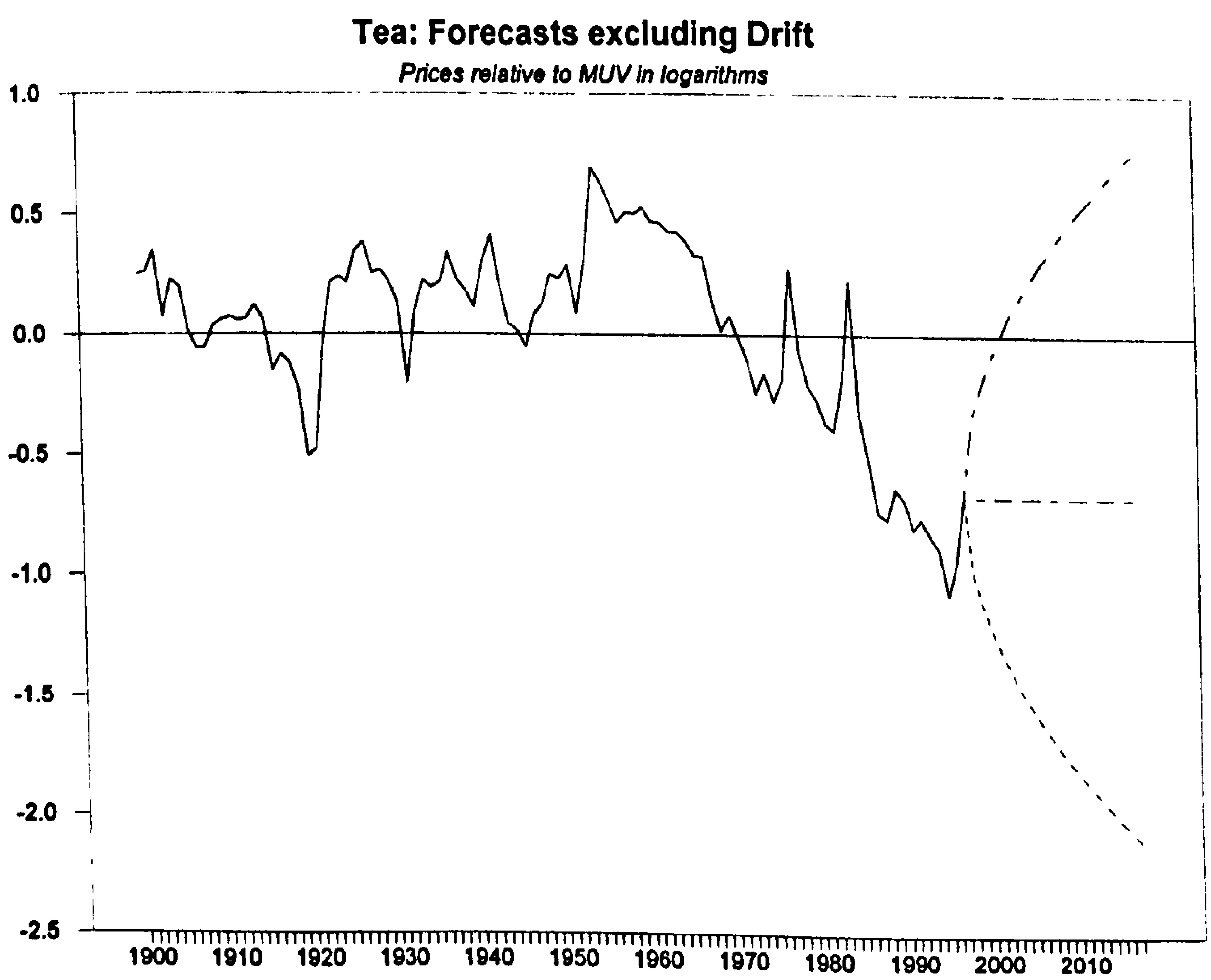


Figure 5.3.10:

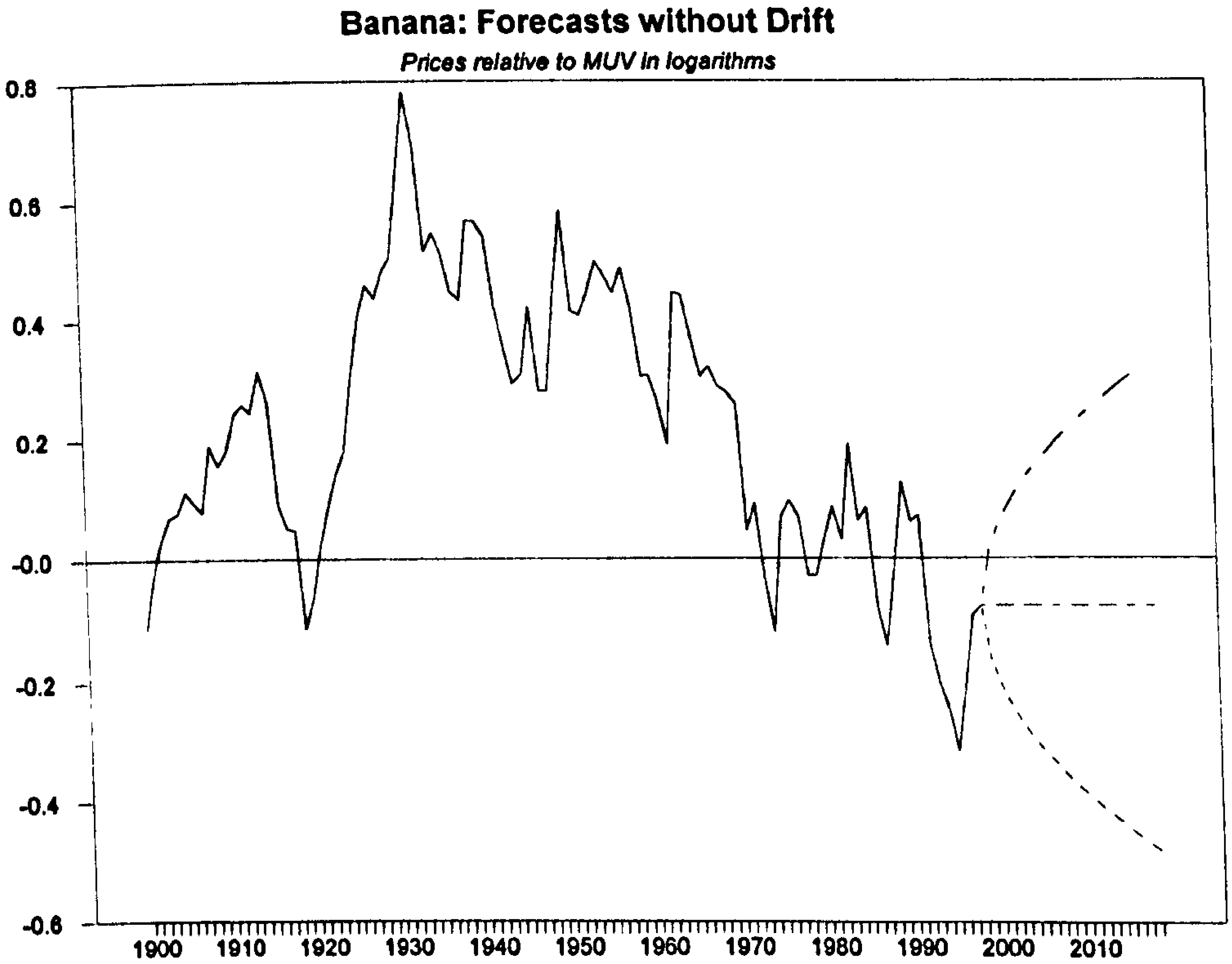


Figure 5.3.11:

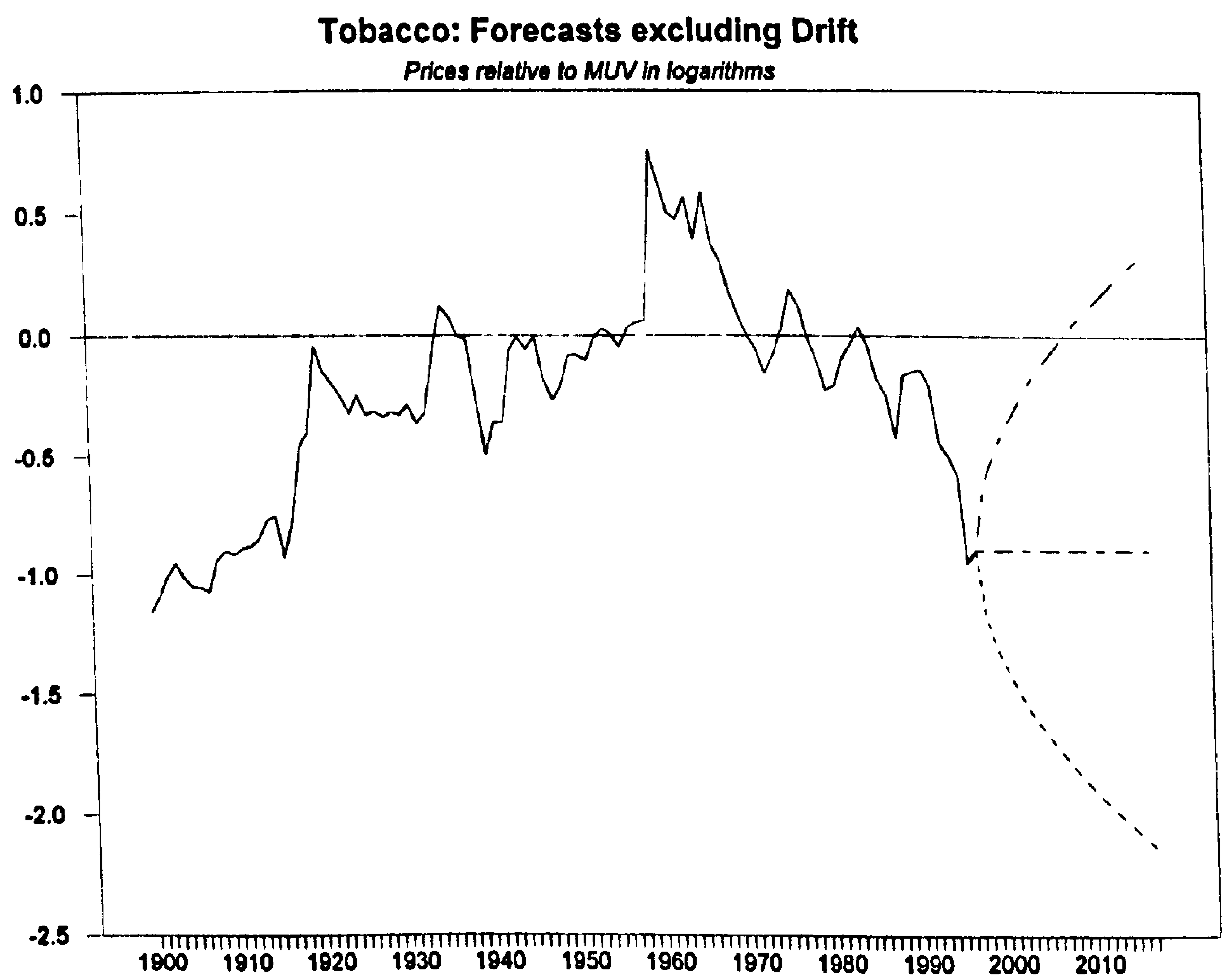


Figure 5.3.12:

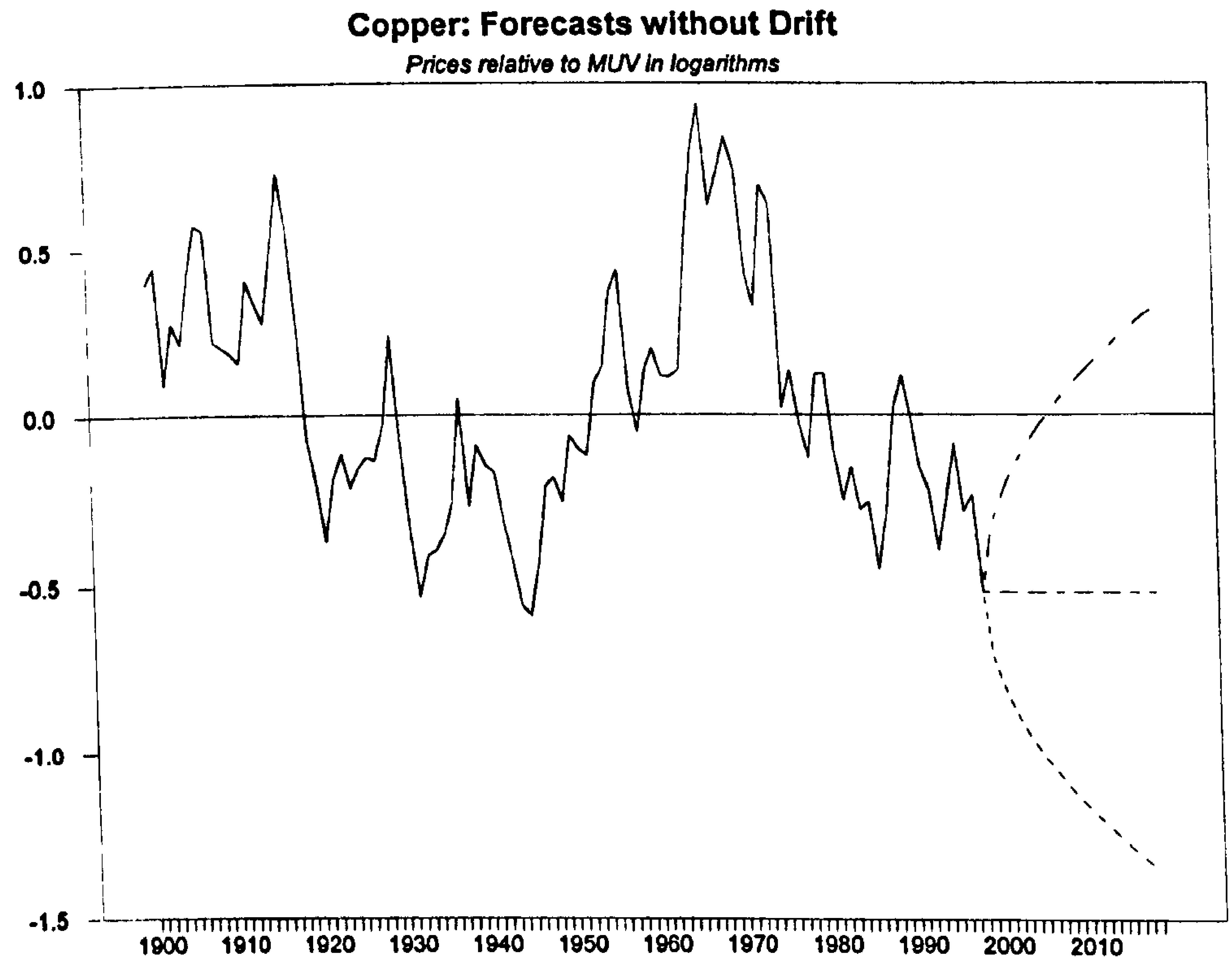
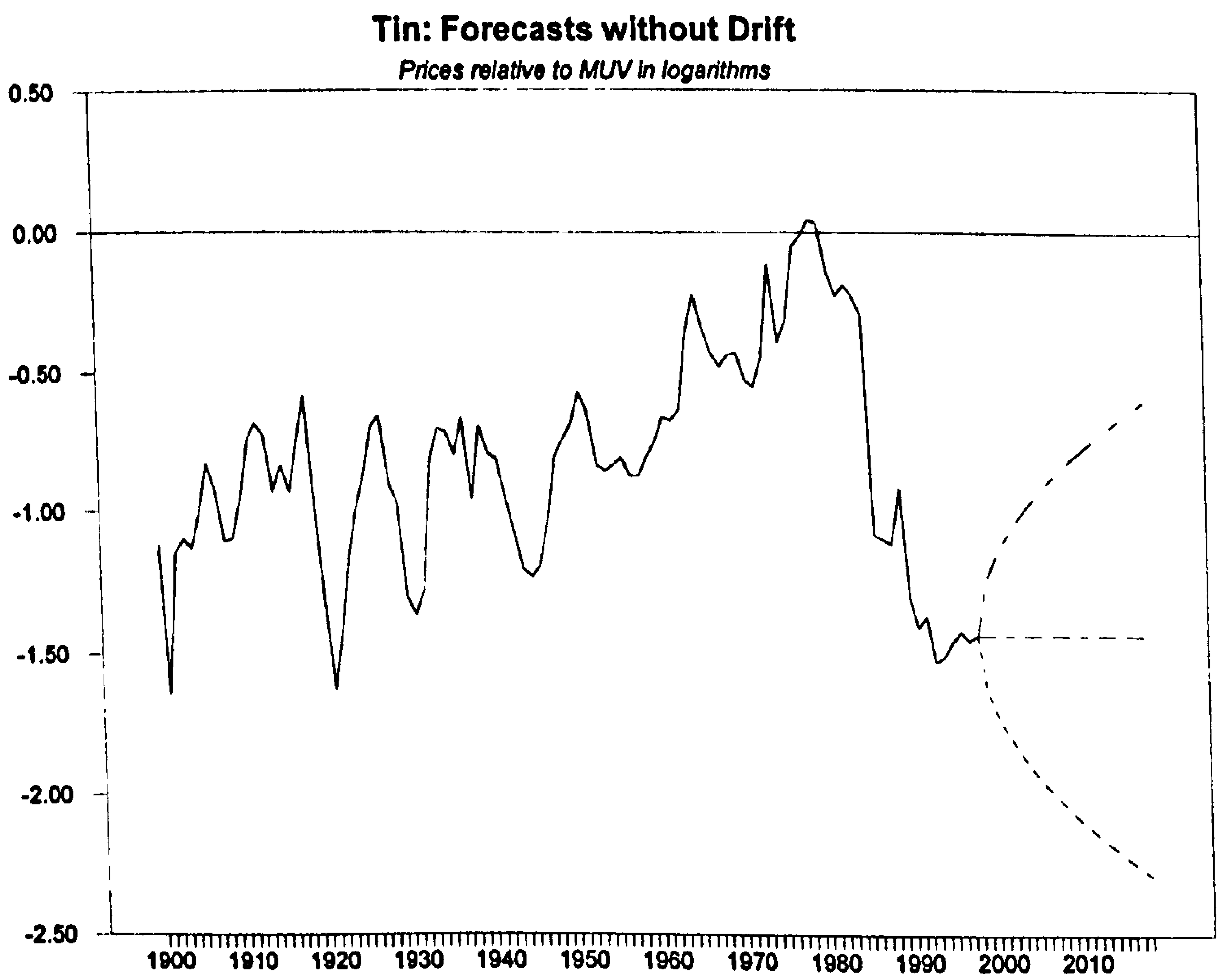


Figure 5.3.13:



It was pointed out above in Chapter 4 that random walk models can easily be obtained from overdifferenced stationary or trend stationary time series and it was further demonstrated that fitting ARIMA(1,1,1) models can yield improved results over pure random walks -at least in the case where the trend stationary alternative would be an ARIMA(1,0,0) model. In the present case forecasts for ARIMA(1,1,1) models without drift are presented in addition to the pure random walk forecasts shown above. Forecasts on the basis of ARIMA(1,1,1) models are shown in table 5.3.5. below.

Table 5.3.5. Point forecasts from ARIMA(1,1,1) models as an alternative to pure random walks

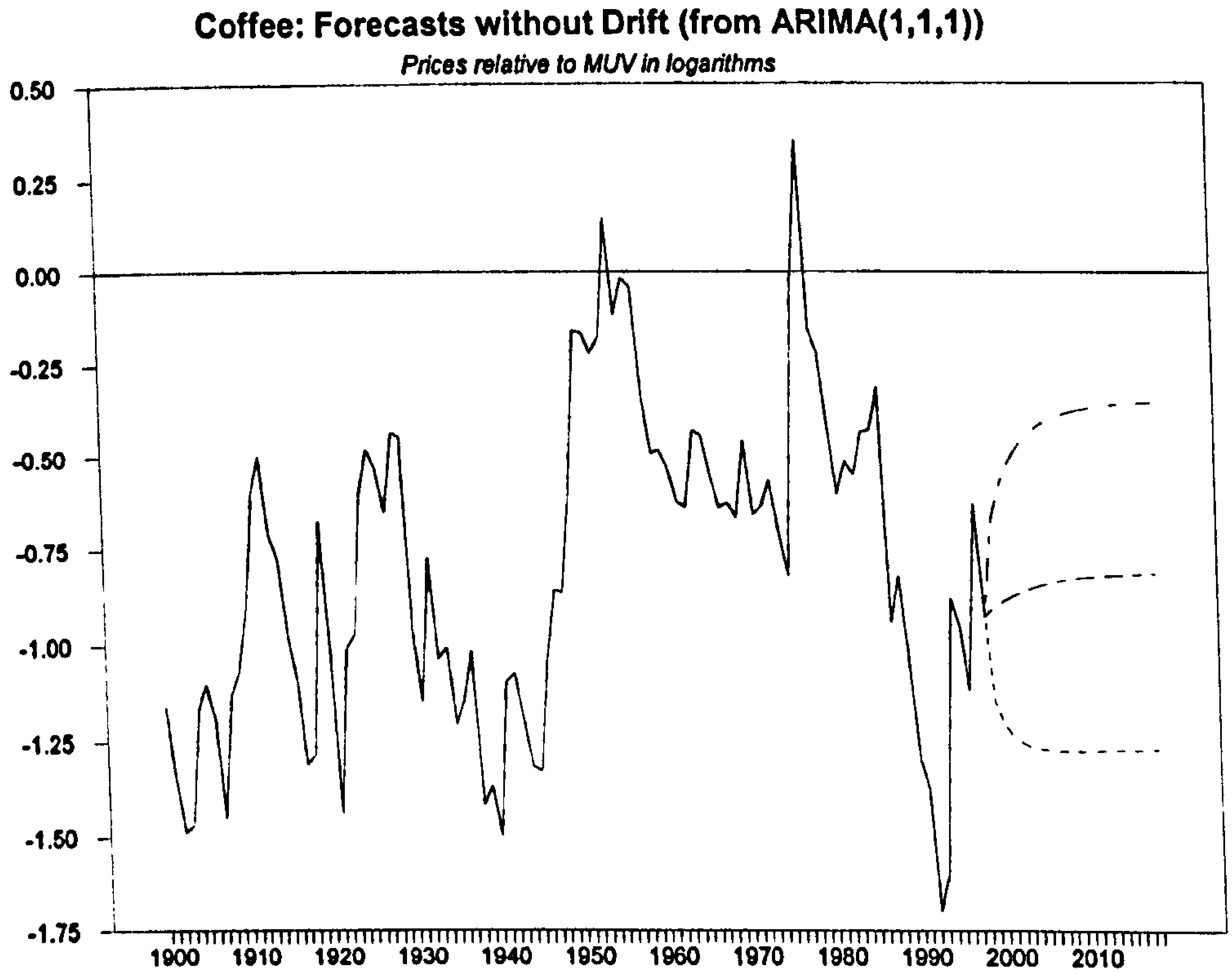
h	Coffee	Tea	Bananas	Tobacco	Copper	Tin
0	*-0.926*	*-0.672*	*-0.077*	*-0.898*	*-0.522*	*-1.441*
1	-0.907	-0.925	-0.079	-0.980	-0.560	-1.439
2	-0.891	-0.946	-0.078	-0.985	-0.545	-1.439
3	-0.877	-0.932	-0.078	-0.986	-0.551	-1.439
4	-0.866	-0.941	-0.078	-0.986	-0.548	-1.439
5	-0.856	-0.935	-0.078	-0.986	-0.549	-1.439
6	-0.848	-0.939	-0.078	-0.986	-0.549	-1.439
7	-0.842	-0.936	-0.078	-0.986	-0.549	-1.439
8	-0.836	-0.938	-0.078	-0.986	-0.549	-1.439
9	-0.831	-0.937	-0.078	-0.986	-0.549	-1.439
10	-0.828	-0.938	-0.078	-0.986	-0.549	-1.439

h: Forecast horizon, *h*=0 last observation of the original data set. All forecasts are for prices relative to MUV in natural logarithms. Forecast models for Tobacco and Tea have been computed on the basis of data up to 1997 only.

For the alternative ARIMA(1,1,1) models point forecasts vary in the short run and then converge to a constant long term forecast value, which in the cases of Tea and Tobacco is quite distinct from the random walk forecast. In the case of Coffee, where the estimated autoregressive coefficient ($\hat{\phi} = 0.931$) is close to one, the convergence process is comparatively slow.

It will be recalled that the ARIMA(1,1,1) model was only selected as a forecast model for Coffee and Lead, where in both cases the autoregressive coefficient is significant while the moving average coefficient estimate is on the invertibility boundary. The forecasts for Coffee from an ARIMA(1,1,1) model over a 20 year period are shown in figure 5.3.14.

Figure 5.3.14:



In all of the other cases listed in table 5.3.5, neither the coefficient estimate on the autoregressive term nor on the moving average term appear statistically significant on the basis of a standard t-test (*cf.* the results reported in Appendix IV.i.). (The one exception is Tea, when the model is re-estimated without drift *cf.* Appendix

V.iii.). Thus a pure random walk is the preferred forecast model for all the remaining commodity price series covered in tables 5.3.5. and 5.3.4.

Ten period forecasts for those commodity price series where there was substantial uncertainty surrounding *a priori* conclusions on the presence of a trend or drift coefficient are given in tables 5.3.6a and 5.3.6.b. below. Again the forecasts in the table are only for those cases in this group where the most appropriate forecast model is not a pure random walk. A random walk was chosen as a forecast model for Beef. The point forecast from a random walk for Beef is -0.816, and would be -1.033 for Lead. The remaining forecasts are:

Table 5.3.6.a Point forecasts for commodity price series where the presence of a trend or drift component is uncertain

h	**Rice** ARIMA(1,1,1)	**Wheat** ARIMA(0,1,4)	**Maize** ARIMA(0,1,2)	**Palm Oil** ARIMA(2,1,0)
0	*-0.583*	*-0.618*	*-0.602*	*-0.463*
1	-0.541	-0.586	-0.506	-0.480
2	-0.561	-0.599	-0.502	-0.568
3	-0.577	-0.534	-0.502	-0.567
4	-0.592	-0.512	-0.502	-0.535
5	-0.605	-0.522	-0.502	-0.534
6	-0.617	-0.533	-0.502	-0.545
7	-0.630	-0.544	-0.502	-0.546
8	-0.642	-0.555	-0.502	-0.542
9	-0.654	-0.565	-0.502	-0.542
10	-0.666	-0.576	-0.502	-0.543

h: Forecast horizon, *h*=0 last observation of the original data set, All forecasts are for prices relative to MUV in natural logarithms. Forecasts for series best modelled by a pure random walk are not included in the table., Viii. Forecast models included a constant in the cases of Rice and Wheat.

Table 5.3.6.b Point forecasts for commodity price series where the presence of a trend or drift component is uncertain

h	**Cotton** ARIMA(2,1,2)	**Wool** ARIMA(0,1,2)	**Lamb** ARIMA(5,0,0)	**Lead** ARIMA(1,1,1)
0	*-0.689*	*-0.738*	*-0.306*	*-1.033*
1	-0.631	-0.744	-0.308	-0.988
2	-0.570	-0.662	-0.126	-0.955
3	-0.534	-0.662	-0.050	-0.931
4	-0.534	-0.662	-0.073	-0.913
5	-0.561	-0.662	-0.017	-0.901
6	-0.596	-0.662	0.099	-0.891
7	-0.621	-0.662	0.154	-0.885
8	-0.628	-0.662	0.168	-0.880
9	-0.618	-0.662	0.213	-0.876
10	-0.599	-0.662	0.273	-0.874

h: Forecast horizon, *h*=0 last observation of the original data set, All forecasts are for prices relative to MUV in natural logarithms. Forecasts for series best modelled by a pure random walk are not included in the table. The forecast model included a trend in the case of Lamb.

The forecasts for Rice and Wheat, show the impact of the downward drift which has been incorporated into the forecast model. In the case of Lamb the impact of the positive trend coefficient estimate can be observed. For the remaining driftless difference stationary forecast models the point forecast converges to constant values which, except for the case of Palm Oil, lie above the value of the last observation in the sample. The convergence to the constant long run forecast is rather slow for Palm Oil, Cotton and Lead. In the case of Palm Oil the predicted series converges after 12 forecast periods (to a value of -0.543) while for Cotton the forecast horizon required for forecasts to converge to a value of -0.593 is of 53 periods, and the predictions for Lead converge onto a constant value of -0.867 after 27 forecast periods. The forecasts in Table 5.3.6 are illustrated in Figures 5.3.14 to 5.3.21 below.

Figure 5.3.14:

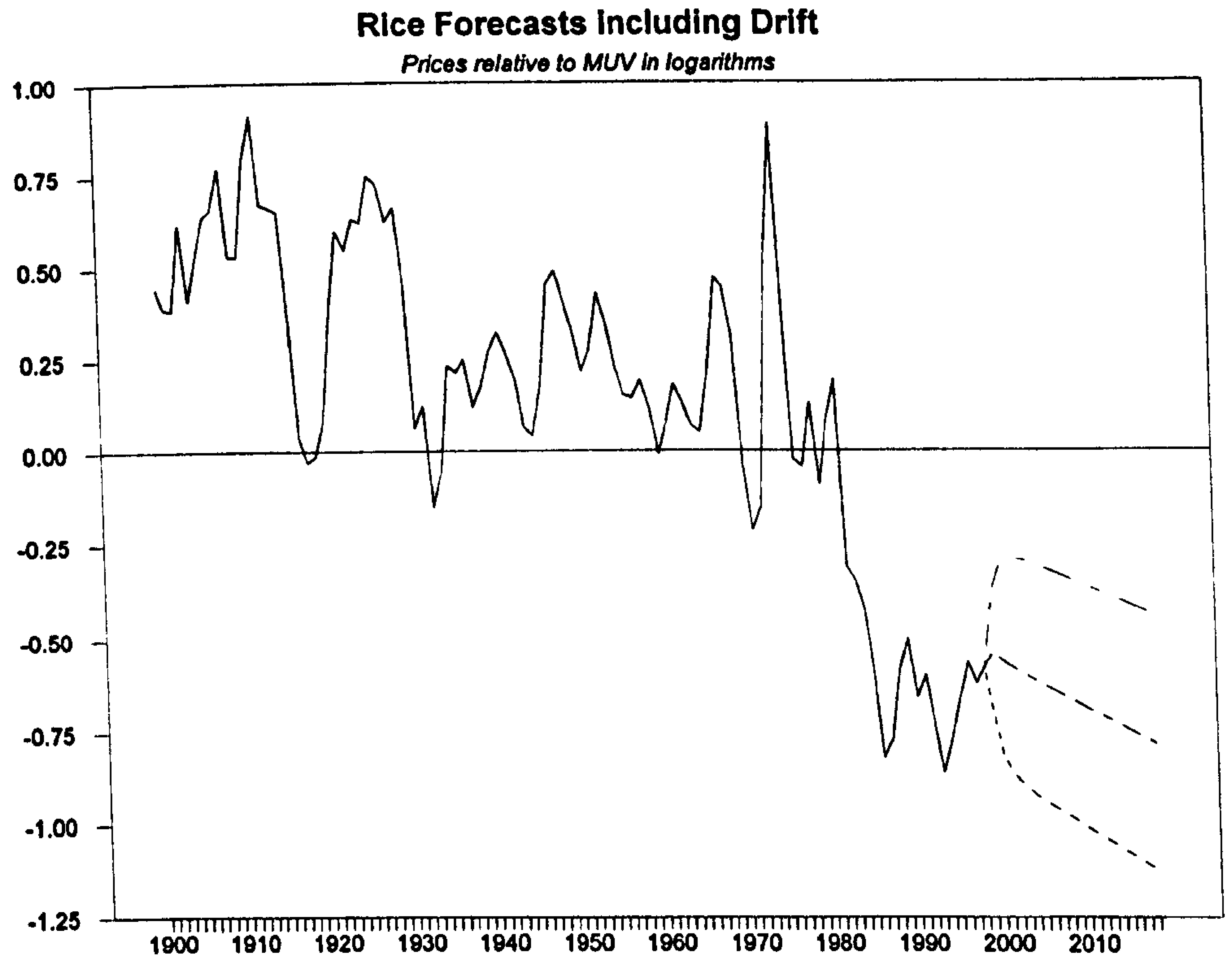


Figure 5.3.15:

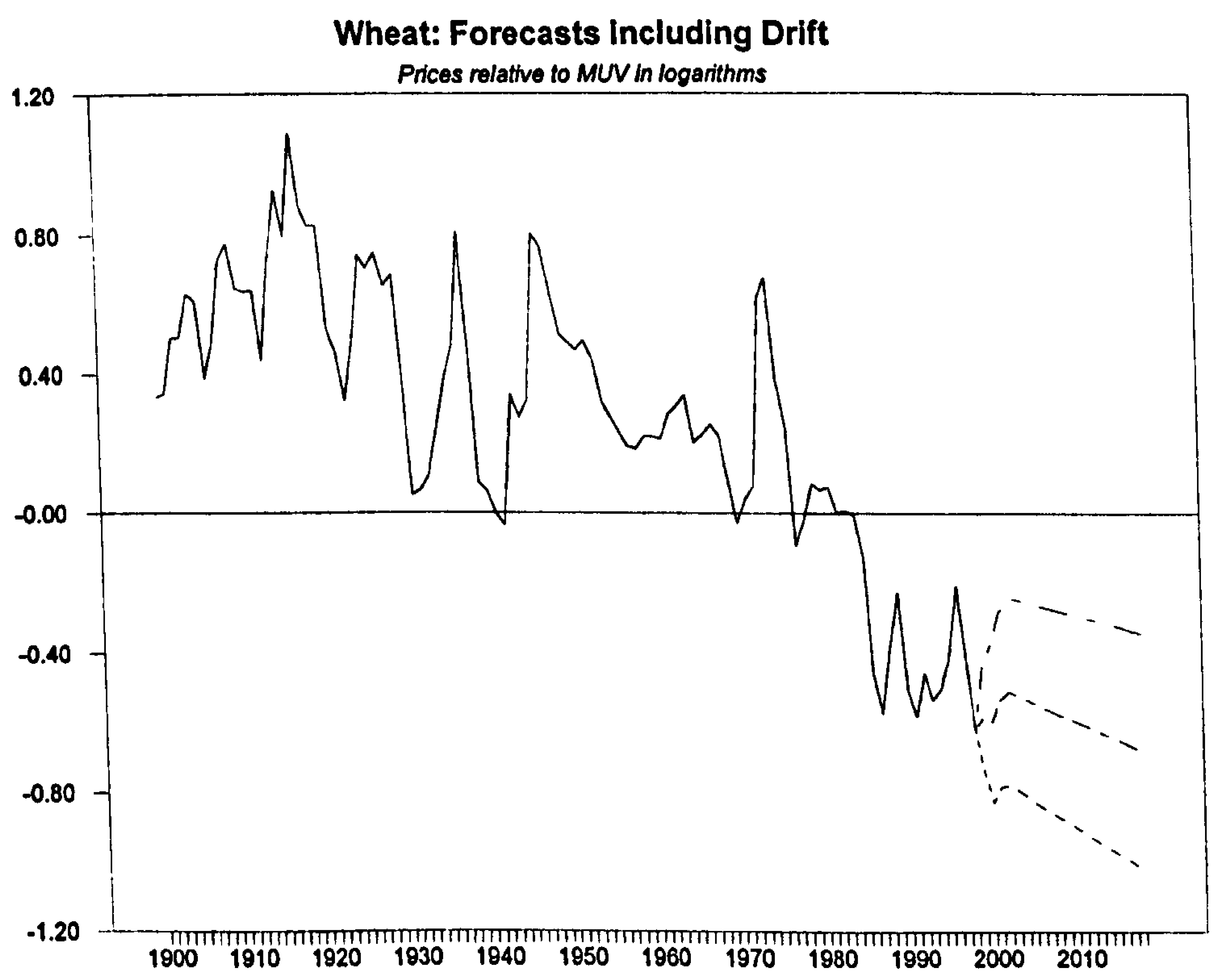


Figure 5.3.16:

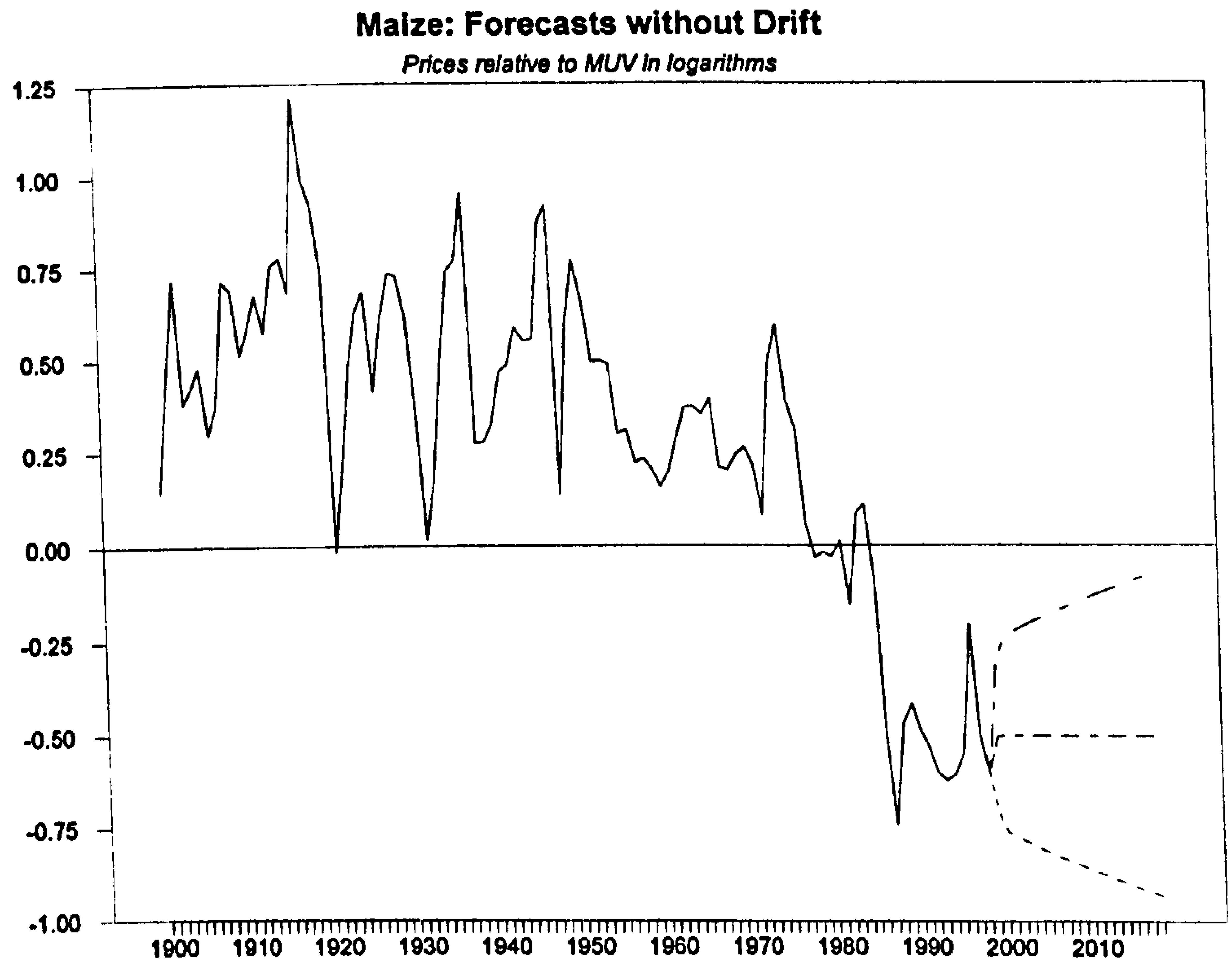


Figure 5.3.17

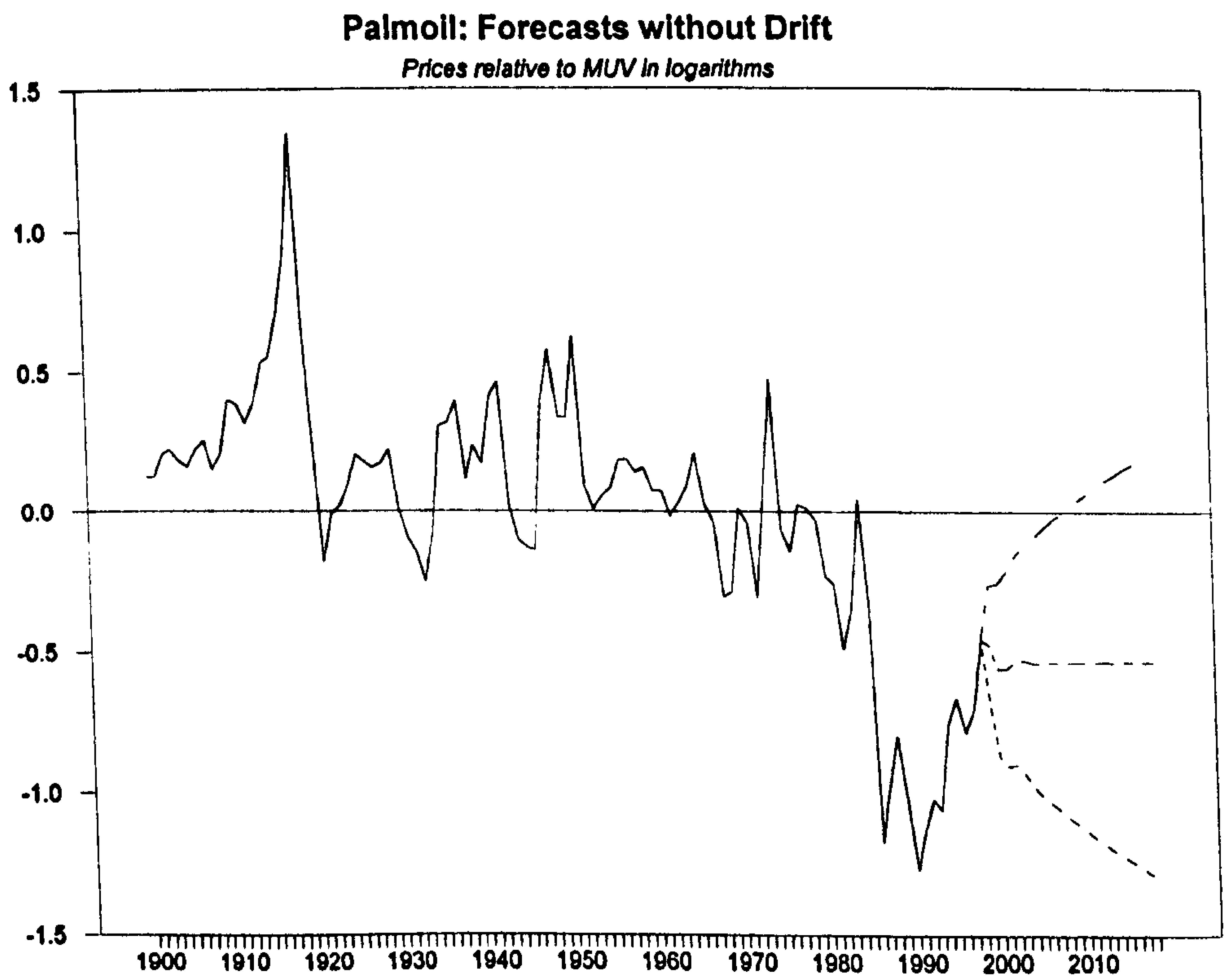


Figure 5.3.18:

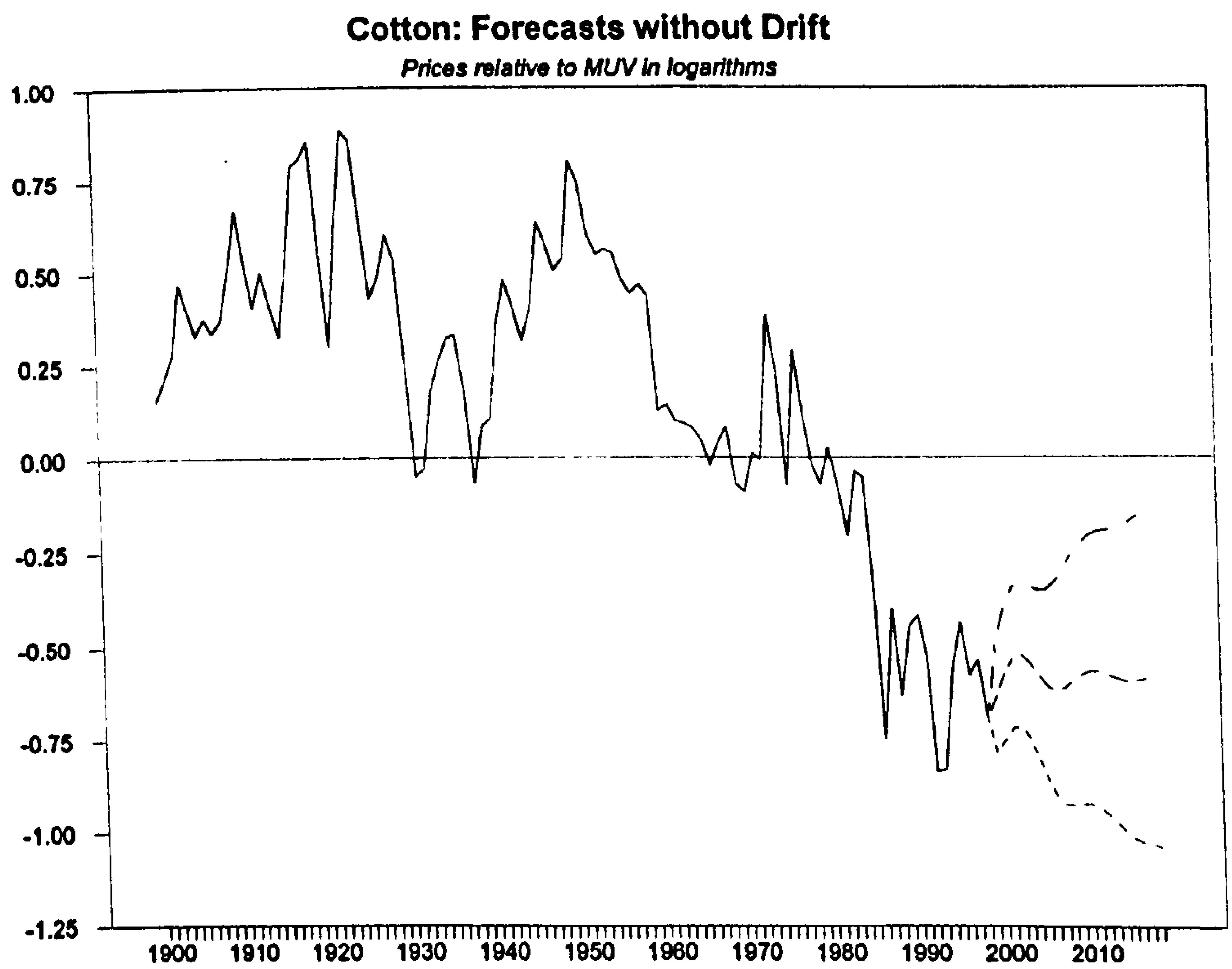


Figure 5.3.19:

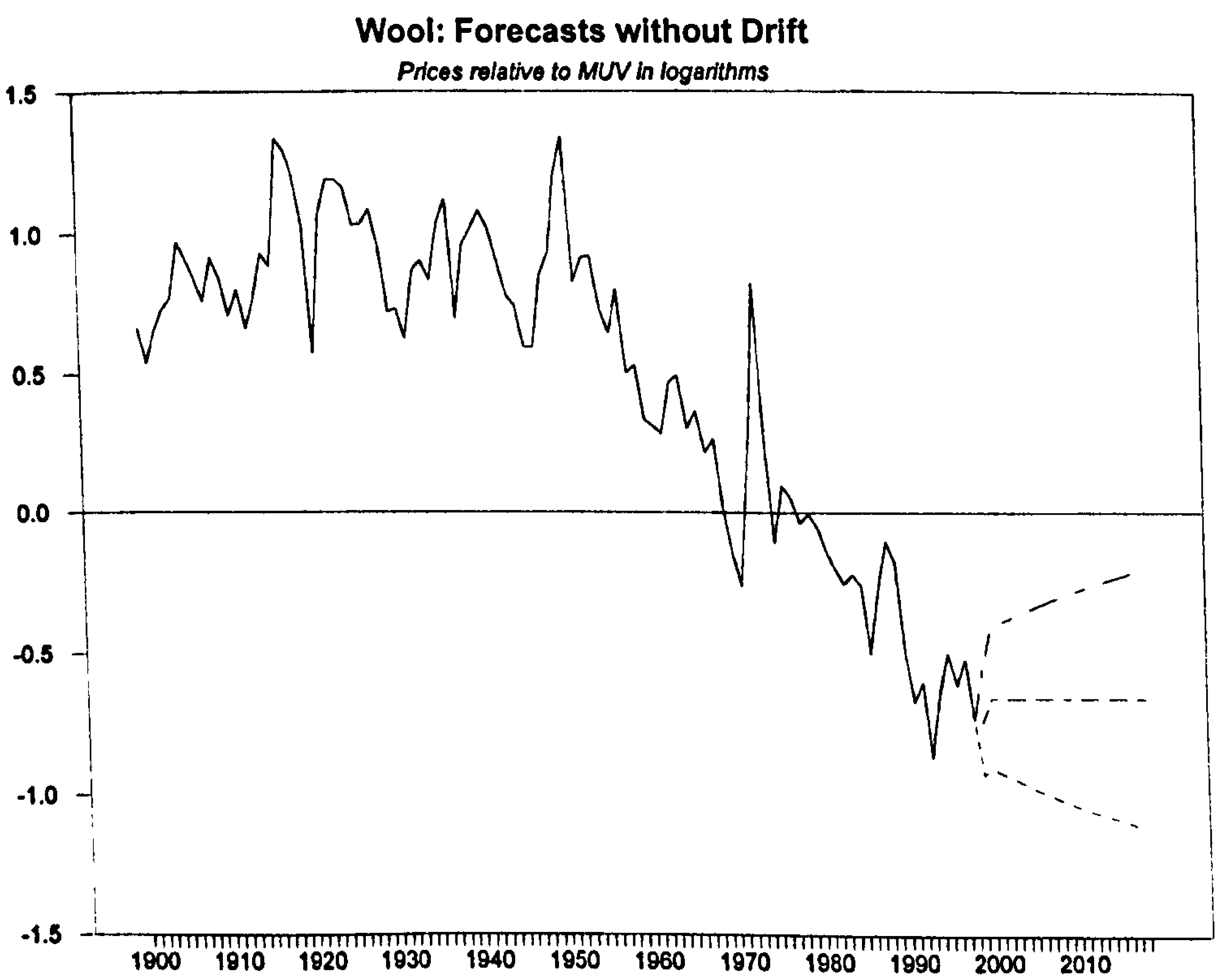


Figure 5.3.20:

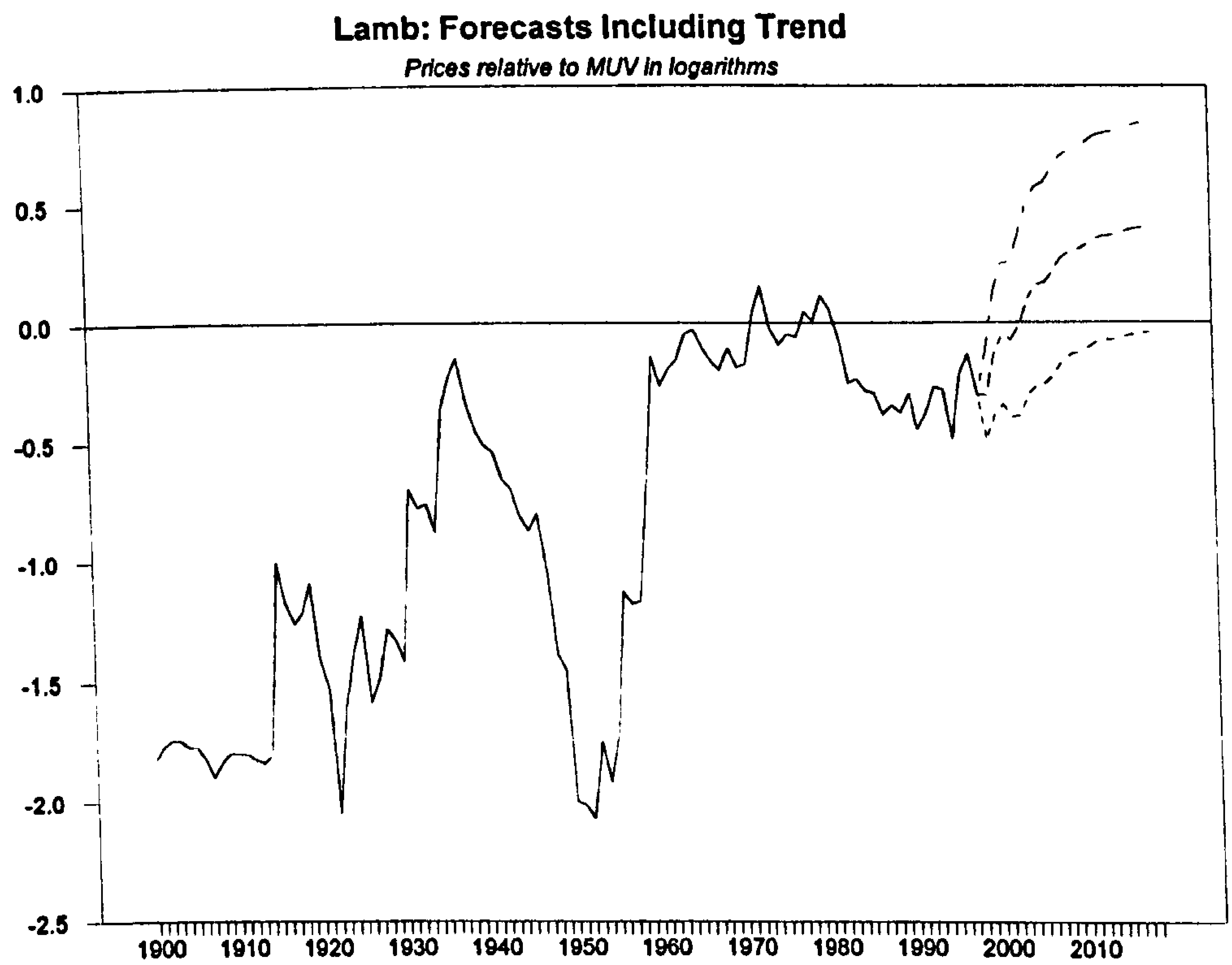
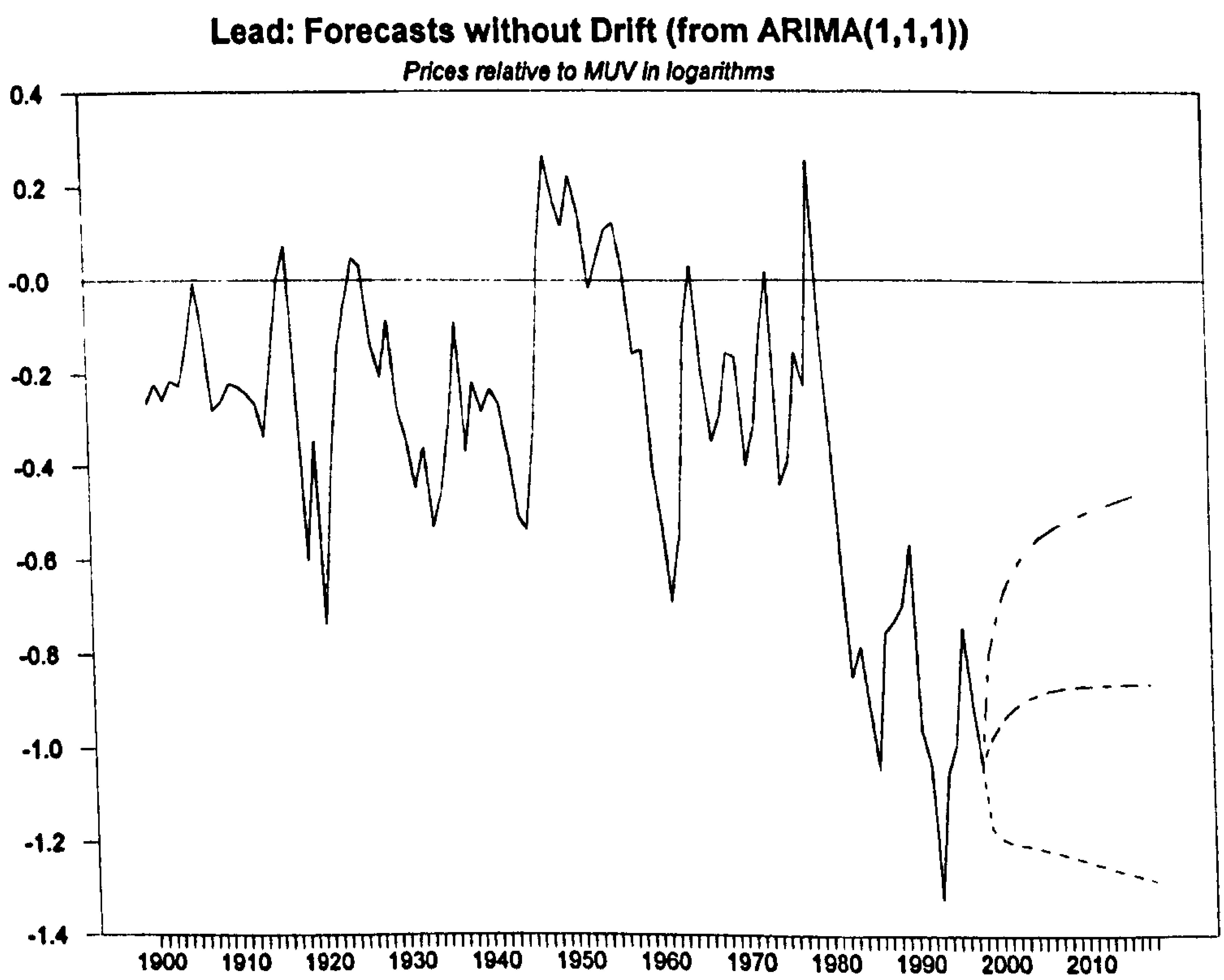


Figure 5.3.21:



The forecasts for Beef, where a pure random walk has been identified as the most appropriate forecast model are based on the model selected by SBC, without drift, yielding a forecast of -0.816. For the reasons given above, these can be compared with forecasts on the basis of ARIMA(1,1,1) models without drift as shown in table 5.3.7. below.

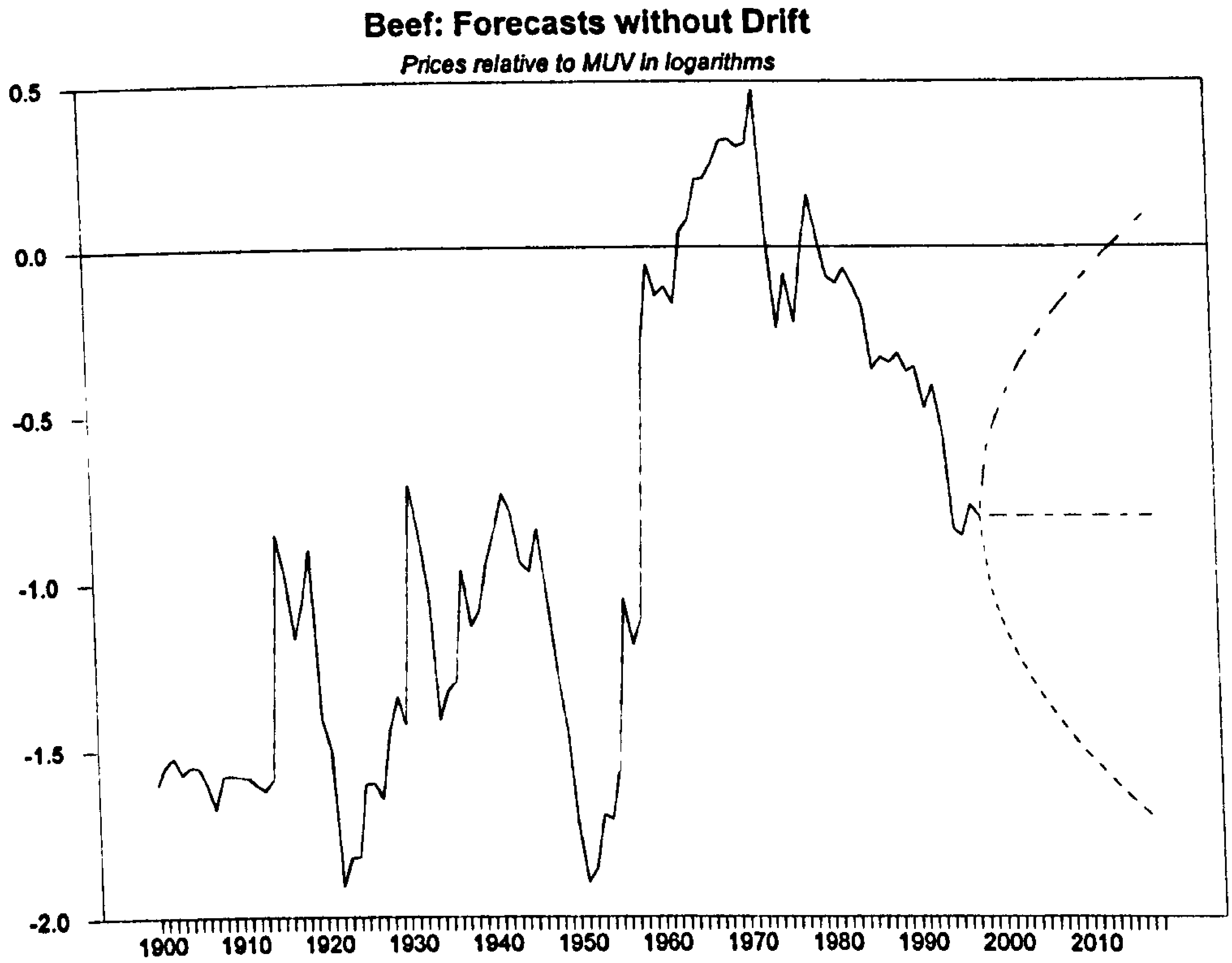
Table 5.3.7. Point forecasts from ARIMA (1,1,1) model for Beef as an alternative to a pure random walk

h	0	1	2	3	4	5
Beef	*-0.816*	-0.820	-0.819	-0.819	-0.819	-0.819
h	6	7	8	9	10	
Beef	-0.819	-0.819	-0.819	-0.819	-0.819	

h: Forecast horizon, *h*=0 last observation of the original data set. All forecasts are for prices relative to MUV in natural logarithms.

As above, the point forecasts converge to a constant long run value distinct from the value predicted by the pure random walk model, after some short run variation is predicted from the ARMA components, although the difference is moderate in this case. For Beef, neither the autoregressive nor moving average coefficient estimates appeared significant in the ARIMA (1,1,1) model, and a pure random walk had been selected earlier as the most appropriate forecast model. The forecasts over 20 years shown in Figure 5.3.22 below are those obtained when modelling the price series for Beef as a pure random walk.

Figure 5.3.22:



5.3.2. Comparison with Worldbank forecasts

The forecasts obtained here can be compared with updates of the data set used as published by the Worldbank (see bibliography for data sources). Commodity price updates and projections where re-indexed to their 1977-1979 average as in the present case and expressed relative to the MUV index, indexed to the same base period and expressed in natural logarithms[20]. Table 5.3.8. shows the projections

[20]Expressing the commodity price series in either constant 1990 US$ or current dollar terms is inappropriate in so far as the interest has been in the development of relative prices throughout. Expressing the price series in constant dollar terms would make it necessary to take the realisation of the of the MUV index as given and interpret the forecasts obtained as simple price projections when what has been predicted are relative commodity price series. Reflation to current dollar terms would moreover be complicated by the fact that reflation of the Worldbank data would have to be accomplished on the basis of a price index which is not publicly available.

from the current study as well as updates and projections of the relevant price series as published by the Worldbank for the years 1999, 2000 and 2001.

Table 5.3.8. Commodity price forecasts for 1999 to 2001 in comparison with World Bank forecasts

Commodity	1999	2000	2001	*1999WB*	*2000WB*	*2001WB*
Coffee	-0.907	-0.891	-0.877	*-1.189*	*-1.390*	*-1.683*
Cocoa	-1.363	-1.389	-1.378	*-1.709*	*-1.959*	*-1.943*
Tea†	-0.672	-0.672	-0.672	*n.a.*	*n.a.*	*n.a.*
Rice	-0.541	-0.561	-0.577	*-0.785*	*-1.015*	*-1.132*
Wheat	-0.586	-0.599	-0.534	*-0.737*	*-0.743*	*-0.633*
Maize	-0.506	-0.502	-0.502	*-0.761*	*n.a.*	*n.a.*
Sugar	-0.290	-0.301	-0.311	*-0.893*	*-0.651*	*-0.465*
Beef	-0.816	-0.816	-0.816	*-0.751*	*-0.728*	*-0.752*
Lamb	-0.308	-0.126	-0.050	*-0.358*	*-0.380*	*-0.331*
Bananas	-0.077	-0.077	-0.077	*-0.351*	*-0.249*	*0.091*
Palm Oil	-0.480	-0.568	-0.567	*-0.895*	*-1.260*	*-1.512*
Cotton	-0.631	-0.570	-0.534	*n.a.*	*n.a.*	*n.a.*
Jute	-0.813	-0.795	-0.795	*-0.873*	*-0.894*	*-0.861*
Wool	-0.744	-0.662	-0.662	*-0.568*	*-0.501*	*-0.496*
Tobacco†	-0.898	-0.898	-0.898	*-0.340*	*n.a.*	*n.a.*
Rubber	-0.826	-0.854	-0.882	*-1.069*	*-0.999*	*-1.099*
Timber	-0.142	-0.043	0.028	*-0.140*	*-0.149*	*-0.270*
Copper	-0.522	-0.522	-0.522	*-0.527*	*-0.455*	*-0.500*
Aluminium	-0.508	-0.529	-0.550	*-0.392*	*-0.287*	*-0.273*
Tin	-1.441	-1.441	-1.441	*-1.465*	*-1.484*	*-1.564*
Silver	-0.817	-0.866	-0.871	*-0.878*	*-0.951*	*-1.051*
Lead	-0.988	-0.955	-0.931	*-1.084*	*-1.211*	*-0.046*
Zinc	-0.131	-0.045	-0.008	*-0.068*	*-0.046*	*-0.164*

† The forecast model for these commodities was estimated on the basis of data from 1900 to 1997 only. All forecasts are for primary commodity prices relative to MUV in natural logarithms.
WB: Worldbank forecasts are given in italics. *n.a.*: not available.

One major complication in comparing the forecast results listed in table 5.3.8., aside from a lack of information about the forecast model specification used for the Worldbank forecasts, is the short forecast horizon considered. At very short forecast horizons, one may expect substantial forecast improvements from the consideration of information from structural models as well as *ad hoc* information of limited availability. The univariate forecasts developed here may prove most valuable at medium to long term horizons.

Generally speaking, there is no general pattern of correspondence or non-correspondence between the forecasts developed here and those of the Worldbank. The point forecasts published by the Worldbank do at times lie within the 68% confidence intervals of the point forecasts from the present study while on other occasions they do not[21].

A comparison in terms of revealed trend movements is equally difficult since the impact of autoregressive and moving average terms can easily dominate the forecast results at the forecast horizon considered for Table 5.3.8. This is for example the case for Wheat, where convergence to the long term negative trend appears to suggest an upwards movement during the three forecast periods shown.

In addition to those commodities for which specific price movements are suggested by the forecasts, no such prediction is made from random walk models. If a random walk model, or a driftless I(1) model more generally, provides a good basis for commodity price forecasts, one should expect the actual future realisations of the data series to lie on either side of the forecast value, rather than corresponding exactly to it. If the Worldbank forecast data are viewed as a benchmark case representative of future realisations of the data series[22], it would still be

[21]The Worldbank's forecasts that lie within the 68% confidence interval given here in the cases of Cocoa, Wheat, Beef, Lamb, Jute, Wool, Rubber, Copper, Tin, Silver and Zinc throughout. Furthermore in the case of Aluminium, Maize and Palm Oil for 1999; Coffee for 2000, Sugar for 2000 and 2001 as well as Timber and Lead for 1999 to 2000. The Worldbank forecasts for Bananas and Rice lie outside the 68% confidence interval for the forecasts obtained here throughout.

[22]At least the data for 1999, and 2000, should be expected to be close to future consolidated figures for these series.

unreasonable to expect a pattern of observations corresponding to a random walk to emerge over the course of just three observations.

5.3.3. Assessing in-sample forecasts.

The quality of the forecast models obtained can be further assessed against the background of the observed discrepancy between in-sample forecasts from models selected in the present chapter and the actual realisations of the data series over the period in question.

For the purpose of this quality assessment, the selected forecast models were re-estimated over the shorter 1900-1983 sample period previously covered by Cuddington (1992). The forecast models themselves are those specified in Table 5.3.1. The models have been re-estimated for the shorter sample period, but since the interest here is in the potential of the forecast model as a characterisation of the long term properties of the series in question, no re-selection has been undertaken.

Comparing the data for the 1984-1994 period with the confidence intervals for the 10 period in-sample forecasts, only a small number of models seems to perform reasonably well. The actual data series lie within the interval forecasts for Rubber, Aluminium and Copper for the entire forecast horizon and take values beyond the confidence intervals in three or less periods in the cases of Jute, Beef, Timber and Lamb.

The data for Palm Oil, Lead and Zinc lie outside the confidence interval at short horizons but then converge to it after 3 to 7 periods. The model for Wool performs well initially but the data series then diverge from the confidence interval after 6 periods. The data series for the remaining commodities lie outside the forecast confidence intervals during four or more periods, and for the entire forecast horizon in the cases of Tea, Cotton, Wheat, Maize, Rice, Cocoa, Silver and Tin.

It is important to be aware of the limitations of such a general comparison with the interval predictions obtained. Confidence interval width differs across forecast models, and is constant only in stationary models while prediction intervals are set to widen where the model selected is difference stationary. Moreover, the forecast performance of many of the selected univariate models improves substantially if the models are re-estimated over the slightly longer 1900-1986 period.

It is also worth noting that most of the trend stationary models perform rather well. It is not clear though that this would also have been so had they not been selected after careful consideration of the evidence on the order of integration. There also appears to be some vindication of the simulation evidence on the role of drift terms in integrated models as the two forecast models including a drift term are among the worst performers.

A further comparison has been made with the results obtained by Cuddington (1992) over the same sample period. The ARIMA model parameterisations were applied as in Cuddington (1992)[23] to obtain forecast models but were re-estimated, since the original data set had been updated from 1960 and beyond.

The forecasts obtained here were compared with those by Cuddington on the basis of the difference between the squared prediction errors from either forecast (with a larger squared error indicating inferior forecast results). The results are summarised in Table 5.3.9. below.

[23] For Silver and Wheat Cuddington identifies two second order Moving Average terms each. In either case the first of these was assumed to be a first order moving average coefficient.

The forecasts obtained in this study have lower prediction errors than those of Cuddington in 8 or more forecast periods for a total of 10 price series, while the opposite is the case for only three cases. In three further cases identical forecasts were obtained, but the remaining results are mixed.

Table 5.3.9. Comparative performance of Forecast models

Commodity	Period or Range	Commodity	Period or Range
Aluminium	$n=8$	Bananas	$n=3$
Rubber	$n=10$	Jute	$h=3$ to 10
Sugar	$n=10$	Tobacco	$h=2$ to 10
Timber	===	Copper	===
Lamb	$h=2$ to 10	Tin	$h=6$ to 10
Rice	$h=2$ to 4	Silver	$n=0$
Wheat	$h=1$ to 6	Maize	$h=2$ to 10
Zinc	$n=4$	Beef	===
Coffee	$h=1$ to 3	Palm Oil	$h=2$ to 10
Cocoa	$h=3$ to 10	Cotton	$n=1$
Tea	$n=0$	Wool	===
Lead	$n=10$		

The Table compares Forecast results on the basis of models selected here with those from the estimates by Cuddington (1992). It is shown for which forecast horizons (h) or number of forecasts (n) Cuddington's models yield inferior forecasts. (=) shows that identical predictions were obtained.

The in-sample evaluation of the forecast results shows that the univariate forecast models, selected partially on the basis of information criteria, can offer a basis for forecasting but should not be relied upon exclusively. The extrapolation of integrated series with drift components seems to require particular care. The observed performance differences brought about by small changes in sample period and forecast horizon also suggest that the pronounced volatility of some data series may itself be one of the main issues to consider.

5.4. Conclusion

It has been shown that projections for relative primary commodity prices should only be based on trend or drift terms in a small number of cases. Where trend or drift terms are statistically significant they may not improve the performance of forecast models unless the point estimate of the trend or drift term is sufficiently large. It was furthermore confirmed that the correct specification of the order of integration of the data generating process underlying the relative price series in question is important for the accuracy of the predictions obtained. While unit root pre-tests can on average lead to improved forecast results in a number of cases, it may be better to rely on differenced forecast models in the case of near integrated series where the power of unit root tests is very low. Conclusions on the usefulness of unit-root pre-testing are altered further, if one explicitly accounts for the interdependence of the assumed order of integration and the evaluation of significance tests on the trend coefficient estimate. In some cases, where unit root pre-testing could be expected to lead to improvements in forecast performance if the presence or absence of a trend component was known in advance, it may still be preferable to rely on a differenced forecast model if the point estimate of the trend coefficient is small and inference on the significance of this estimate depends on the inferred order of integration of the series.

On the basis of the results on forecast model selection presented here, a trend and two drift terms respectively[24] were incorporated in three forecast models where the presence of a trend component has been regarded as uncertain in modelling the

[24] A trend term was used in the forecasts for Lamb, while a drift term was incorporated in the forecast models for Rice and Wheat.

available data series. A trend was further included in the case of those four commodities where its presence had been inferred with a high degree of confidence (*i.e.* Aluminium, Rubber, Sugar and Timber). For the majority of commodity price series, however it would appear, based on the above simulation results, that the best predictions are obtained by forecast models excluding trend or drift terms. In a number of cases, commodity price series are best modelled as pure random walks for forecasting purposes. In any case the overall scenario does not seem to be one of generalised secular downwards trends -neither from a backward looking nor from a forecasting perspective.

Appendix V.i Ten Period Forecasts and Confidence Intervals

This appendix lists ten period ahead forecasts obtained from the forecast models identified in Table 5.3.1 together with the upper and lower bounds of the 68% confidence intervals corresponding to those point forecasts. The lower bound of the confidence interval is identified as *Commodity Name*-L, the upper bound as *Commodity Name*-U and the point forecast as *Commodity Name*-F. As before in Chapter 5, the first data row of each table contains the last observation in the original data series.

One should take account of the fact that the confidence intervals reported below reflect forecast uncertainty on the assumption that the correct model has been used for the forecast; no account is taken of any remaining uncertainty surrounding the specification of the appropriate forecast model.

Year	Alu-L	Alu-F	Alu-U	Rubber-L	Rubber-F	Rubber-U
1998	*-0.394*	*-0.394*	*-0.394*	*-0.931*	*-0.931*	*-0.931*
1999	-0.657	-0.508	-0.359	-1.279	-0.826	-0.374
2000	-0.754	-0.529	-0.305	-1.307	-0.854	-0.402
2001	-0.799	-0.550	-0.300	-1.335	-0.882	-0.429
2002	-0.829	-0.569	-0.310	-1.363	-0.910	-0.457
2003	-0.852	-0.589	-0.326	-1.391	-0.938	-0.485
2004	-0.873	-0.608	-0.343	-1.419	-0.966	-0.513
2005	-0.893	-0.627	-0.361	-1.447	-0.994	-0.541
2006	-0.912	-0.646	-0.380	-1.475	-1.022	-0.569
2007	-0.931	-0.665	-0.399	-1.502	-1.050	-0.597
2008	-0.950	-0.683	-0.417	-1.530	-1.078	-0.625

Year	Sugar-L	Sugar-F	Sugar-U	Timber-L	Timber-F	Timber-U
1998	*-0.540*	*-0.540*	*-0.540*	*-0.282*	*-0.282*	*-0.282*
1999	-0.707	-0.290	0.127	-0.290	-0.142	0.007
2000	-0.718	-0.301	0.116	-0.222	-0.043	0.137
2001	-0.728	-0.311	0.106	-0.164	0.028	0.220

2002	-0.738	-0.321	0.096	-0.118	0.080	0.278
2003	-0.749	-0.332	0.085	-0.082	0.119	0.319
2004	-0.759	-0.342	0.075	-0.053	0.149	0.350
2005	-0.770	-0.353	0.064	-0.029	0.173	0.375
2006	-0.780	-0.363	0.054	-0.010	0.193	0.395
2007	-0.790	-0.373	0.044	0.007	0.210	0.412
2008	-0.801	-0.384	0.033	0.023	0.225	0.428

Year	**Coffee-L**	**Coffee-F**	**Coffee-U**	**Cocoa-L**	**Cocoa-F**	**Cocoa-U**
1998	*-0.926*	*-0.926*	*-0.926*	*-1.319*	*-1.319*	*-1.319*
1999	-1.384	-0.907	-0.430	-1.840	-1.363	-0.886
2000	-1.515	-0.891	-0.266	-2.092	-1.389	-0.686
2001	-1.590	-0.877	-0.165	-2.174	-1.378	-0.582
2002	-1.635	-0.866	-0.096	-2.237	-1.369	-0.500
2003	-1.664	-0.856	-0.048	-2.324	-1.371	-0.419
2004	-1.684	-0.848	-0.013	-2.408	-1.374	-0.340
2005	-1.696	-0.842	0.013	-2.479	-1.374	-0.269
2006	-1.705	-0.836	0.032	-2.543	-1.373	-0.203
2007	-1.710	-0.831	0.047	-2.606	-1.373	-0.140
2008	-1.714	-0.828	0.059	-2.667	-1.373	-0.080

Year	**Bananas-L**	**Bananas-F**	**Bananas-U**	**Jute-L**	**Jute-F**	**Jute-U**
1998	*-0.077*	*-0.077*	*-0.077*	*-0.942*	*-0.942*	*-0.942*
1999	-0.167	-0.077	0.014	-1.232	-0.813	-0.395
2000	-0.204	-0.077	0.051	-1.374	-0.795	-0.216
2001	-0.233	-0.077	0.080	-1.420	-0.795	-0.170
2002	-0.257	-0.077	0.104	-1.463	-0.795	-0.127
2003	-0.279	-0.077	0.125	-1.504	-0.795	-0.087
2004	-0.298	-0.077	0.145	-1.542	-0.795	-0.049
2005	-0.316	-0.077	0.162	-1.578	-0.795	-0.013
2006	-0.332	-0.077	0.179	-1.613	-0.795	0.022
2007	-0.348	-0.077	0.194	-1.646	-0.795	0.055
2008	-0.362	-0.077	0.209	-1.678	-0.795	0.087

Year	**Copper-L**	**Copper-F**	**Copper-U**	**Tin-L**	**Tin-F**	**Tin-U**
1998	*-0.522*	*-0.522*	*-0.522*	*-1.441*	*-1.441*	*-1.441*
1999	-0.709	-0.522	-0.335	-1.629	-1.441	-1.253
2000	-0.787	-0.522	-0.257	-1.707	-1.441	-1.175
2001	-0.846	-0.522	-0.198	-1.766	-1.441	-1.115
2002	-0.896	-0.522	-0.148	-1.817	-1.441	-1.064
2003	-0.941	-0.522	-0.103	-1.861	-1.441	-1.020

2004	-0.981	-0.522	-0.063	-1.901	-1.441	-0.980
2005	-1.017	-0.522	-0.027	-1.938	-1.441	-0.943
2006	-1.052	-0.522	0.007	-1.973	-1.441	-0.909
2007	-1.084	-0.522	0.040	-2.005	-1.441	-0.876
2008	-1.114	-0.522	0.070	-2.036	-1.441	-0.846

Year	Silver-L	Silver-F	Silver-U	Zinc-L	Zinc-F	Zinc-U
1998	*-0.825*	*-0.825*	*-0.825*	*-0.117*	*-0.117*	*-0.117*
1999	-1.169	-0.817	-0.465	-0.325	-0.131	0.063
2000	-1.375	-0.866	-0.357	-0.294	-0.045	0.205
2001	-1.442	-0.871	-0.300	-0.266	-0.008	0.251
2002	-1.480	-0.856	-0.232	-0.252	0.008	0.269
2003	-1.539	-0.854	-0.169	-0.245	0.015	0.276
2004	-1.601	-0.858	-0.116	-0.243	0.018	0.279
2005	-1.652	-0.859	-0.066	-0.241	0.019	0.280
2006	-1.697	-0.858	-0.018	-0.241	0.020	0.281
2007	-1.742	-0.857	0.027	-0.240	0.020	0.281
2008	-1.785	-0.858	0.070	-0.240	0.020	0.281

Year	Rice-L	Rice-F	Rice-U	Wheat-L	Wheat-F	Wheat-U
1998	*-0.583*	*-0.583*	*-0.583*	*-0.618*	*-0.618*	*-0.618*
1999	-0.704	-0.541	-0.378	-0.738	-0.586	-0.434
2000	-0.816	-0.561	-0.306	-0.824	-0.599	-0.374
2001	-0.861	-0.577	-0.293	-0.784	-0.534	-0.284
2002	-0.889	-0.592	-0.295	-0.780	-0.512	-0.244
2003	-0.909	-0.605	-0.301	-0.796	-0.522	-0.249
2004	-0.926	-0.617	-0.309	-0.811	-0.533	-0.255
2005	-0.942	-0.630	-0.317	-0.826	-0.544	-0.261
2006	-0.957	-0.642	-0.326	-0.842	-0.555	-0.268
2007	-0.972	-0.654	-0.335	-0.857	-0.565	-0.274
2008	-0.987	-0.666	-0.345	-0.872	-0.576	-0.280

Year	Maize-L	Maize-F	Maize-U	Beef-L	Beef-F	Beef-U
1998	*-0.602*	*-0.602*	*-0.602*	*-0.816*	*-0.816*	*-0.816*
1999	-0.909	-0.506	-0.104	-1.020	-0.816	-0.612
2000	-1.019	-0.502	0.016	-1.105	-0.816	-0.528
2001	-1.043	-0.502	0.039	-1.169	-0.816	-0.463
2002	-1.066	-0.502	0.062	-1.224	-0.816	-0.408
2003	-1.088	-0.502	0.084	-1.272	-0.816	-0.360
2004	-1.109	-0.502	0.105	-1.316	-0.816	-0.316
2005	-1.129	-0.502	0.125	-1.356	-0.816	-0.276

2006	-1.149	-0.502	0.145	-1.393	-0.816	-0.239
2007	-1.168	-0.502	0.164	-1.428	-0.816	-0.204
2008	-1.186	-0.502	0.183	-1.461	-0.816	-0.171

Year	**Lamb -L**	**Lamb-F**	**Lamb-U**	**Palm Oil-L**	**Palm Oil-F**	**Palm Oil-U**
1998	*-0.306*	*-0.306*	*-0.306*	*-0.463*	*-0.463*	*-0.463*
1999	-0.500	-0.308	-0.116	-0.898	-0.480	-0.061
2000	-0.384	-0.126	0.132	-1.176	-0.568	0.039
2001	-0.348	-0.050	0.248	-1.241	-0.567	0.107
2002	-0.400	-0.073	0.254	-1.264	-0.535	0.193
2003	-0.392	-0.017	0.358	-1.333	-0.534	0.265
2004	-0.303	0.099	0.501	-1.412	-0.545	0.322
2005	-0.262	0.154	0.570	-1.470	-0.546	0.378
2006	-0.256	0.168	0.592	-1.518	-0.542	0.434
2007	-0.218	0.213	0.644	-1.569	-0.542	0.486
2008	-0.160	0.273	0.705	-1.620	-0.543	0.534

Year	**Cotton-L**	**Cotton-F**	**Cotton-U**	**Wool-L**	**Wool-F**	**Wool-U**
1998	*-0.689*	*-0.689*	*-0.689*	*-0.738*	*-0.738*	*-0.738*
1999	-0.923	-0.631	-0.339	-1.112	-0.744	-0.377
2000	-0.919	-0.570	-0.221	-1.147	-0.662	-0.176
2001	-0.903	-0.534	-0.165	-1.179	-0.662	-0.144
2002	-0.918	-0.534	-0.150	-1.209	-0.662	-0.115
2003	-0.969	-0.561	-0.152	-1.237	-0.662	-0.086
2004	-1.047	-0.596	-0.144	-1.264	-0.662	-0.060
2005	-1.132	-0.621	-0.110	-1.289	-0.662	-0.034
2006	-1.201	-0.628	-0.055	-1.314	-0.662	-0.009
2007	-1.242	-0.618	0.007	-1.338	-0.662	0.015
2008	-1.261	-0.599	0.064	-1.361	-0.662	0.038

Year	Lead-L	Lead-F	Lead-U
1998	*-1.033*	*-1.033*	*-1.033*
1999	-1.346	-0.988	-0.629
2000	-1.420	-0.955	-0.490
2001	-1.459	-0.931	-0.403
2002	-1.485	-0.913	-0.342
2003	-1.504	-0.901	-0.297
2004	-1.520	-0.891	-0.263
2005	-1.534	-0.885	-0.235
2006	-1.548	-0.880	-0.212
2007	-1.560	-0.876	-0.192
2008	-1.573	-0.874	-0.175

Year	Tea-L	Tea-F	Tea-U	Tobacco-L	Tobacco-F	Tobacco-U
1997	*-0.672*	*-0.672*	*-0.672*	*-0.898*	*-0.898*	*-0.898*
1998	-0.992	-0.672	-0.353	-1.174	-0.898	-0.622
1999	-1.124	-0.672	-0.220	-1.288	-0.898	-0.508
2000	-1.226	-0.672	-0.119	-1.376	-0.898	-0.420
2001	-1.311	-0.672	-0.033	-1.450	-0.898	-0.347
2002	-1.387	-0.672	0.042	-1.515	-0.898	-0.281
2003	-1.455	-0.672	0.110	-1.574	-0.898	-0.223
2004	-1.518	-0.672	0.173	-1.628	-0.898	-0.168
2005	-1.576	-0.672	0.231	-1.678	-0.898	-0.118
2006	-1.631	-0.672	0.286	-1.725	-0.898	-0.071
2007	-1.683	-0.672	0.338	-1.770	-0.898	-0.026

Appendix V.ii. Forecasts for Wheat Prices from the Difference Stationary Model Selected by SBC

It will be recalled that the model initially selected for Wheat, using minimum SBC was a difference stationary model without drift, and that a drift term was incorporated only after the t-ratio on the trend coefficient had increased substantially when model selection by AIC was considered to correct for overdifferencing. In this Appendix, forecasts obtained from the driftless model selected by SBC are considered as an alternative. The SBC model re-estimated without drift is:

Model: ARIMA(0,1,2) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12)**[1]:21.759 (0.016)

$$\Delta p_t = v_t$$
$$v_t = \underset{(0.159)}{\varepsilon_t} + \underset{(0.086)}{0.135\varepsilon_{t-1}} - \underset{(0.087)}{0.546\varepsilon_{t-2}}$$

The corresponding 10 year forecasts with 68% confidence intervals are given in table V.ii.i. below, where as before the first data row contains the last observation in the original data series.

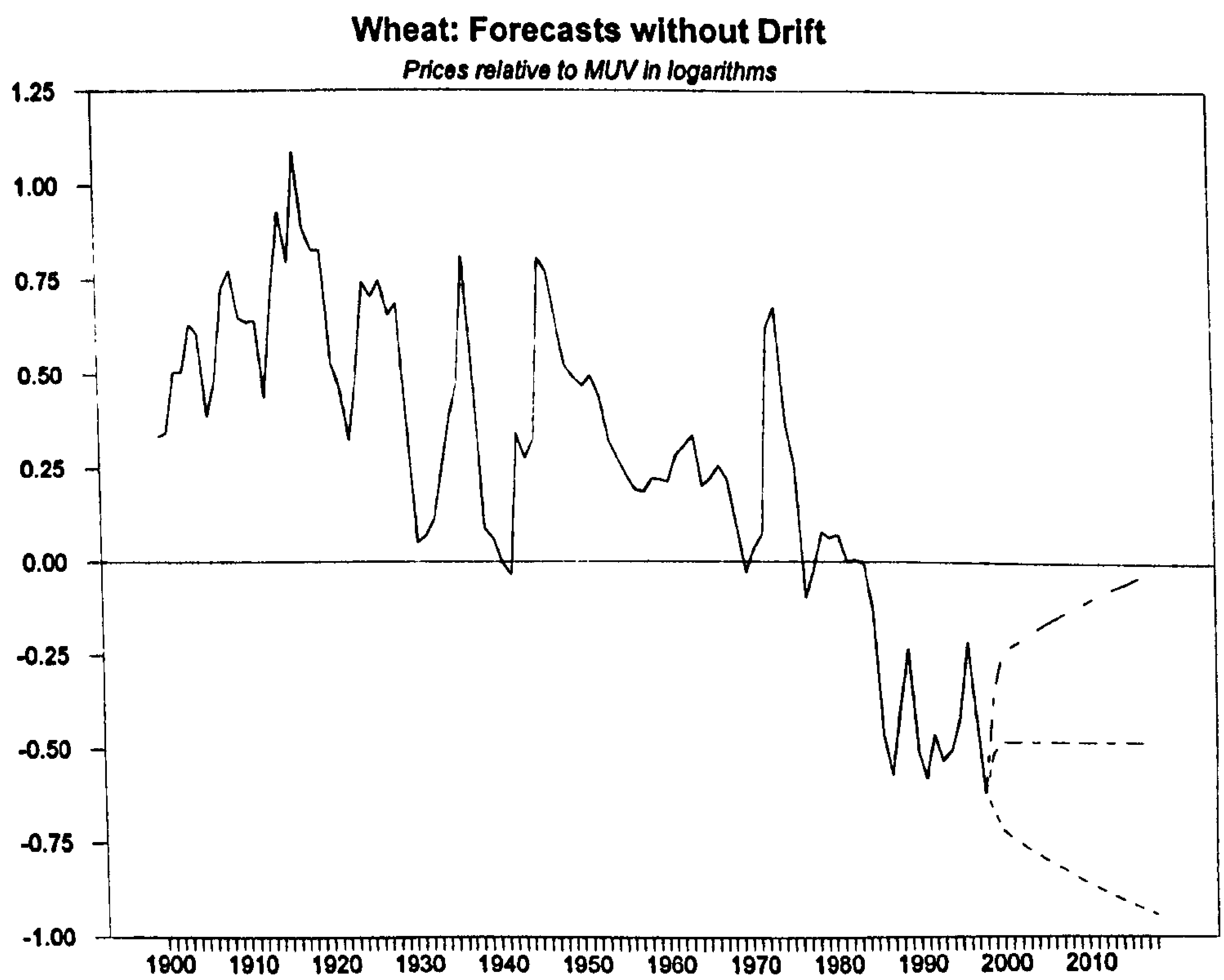
[1] The Ljung Box Q statistic is computed for 12 autocorrelations, the number in parentheses is the associated P-value, indicating a possible problem of serial correlation in the present case.

Table V.ii.i. Ten year forecasts for Wheat based on an ARIMA(0,1,2) model without drift.

Year	Wheat-L	Wheat-F	Wheat-U
1998	*-0.618*	*-0.618*	*-0.618*
1999	-0.822	-0.511	-0.200
2000	-0.952	-0.481	-0.011
2001	-0.986	-0.481	0.023
2002	-1.018	-0.481	0.056
2003	-1.049	-0.481	0.086
2004	-1.078	-0.481	0.115
2005	-1.105	-0.481	0.143
2006	-1.132	-0.481	0.169
2007	-1.157	-0.481	0.195
2008	-1.181	-0.481	0.219

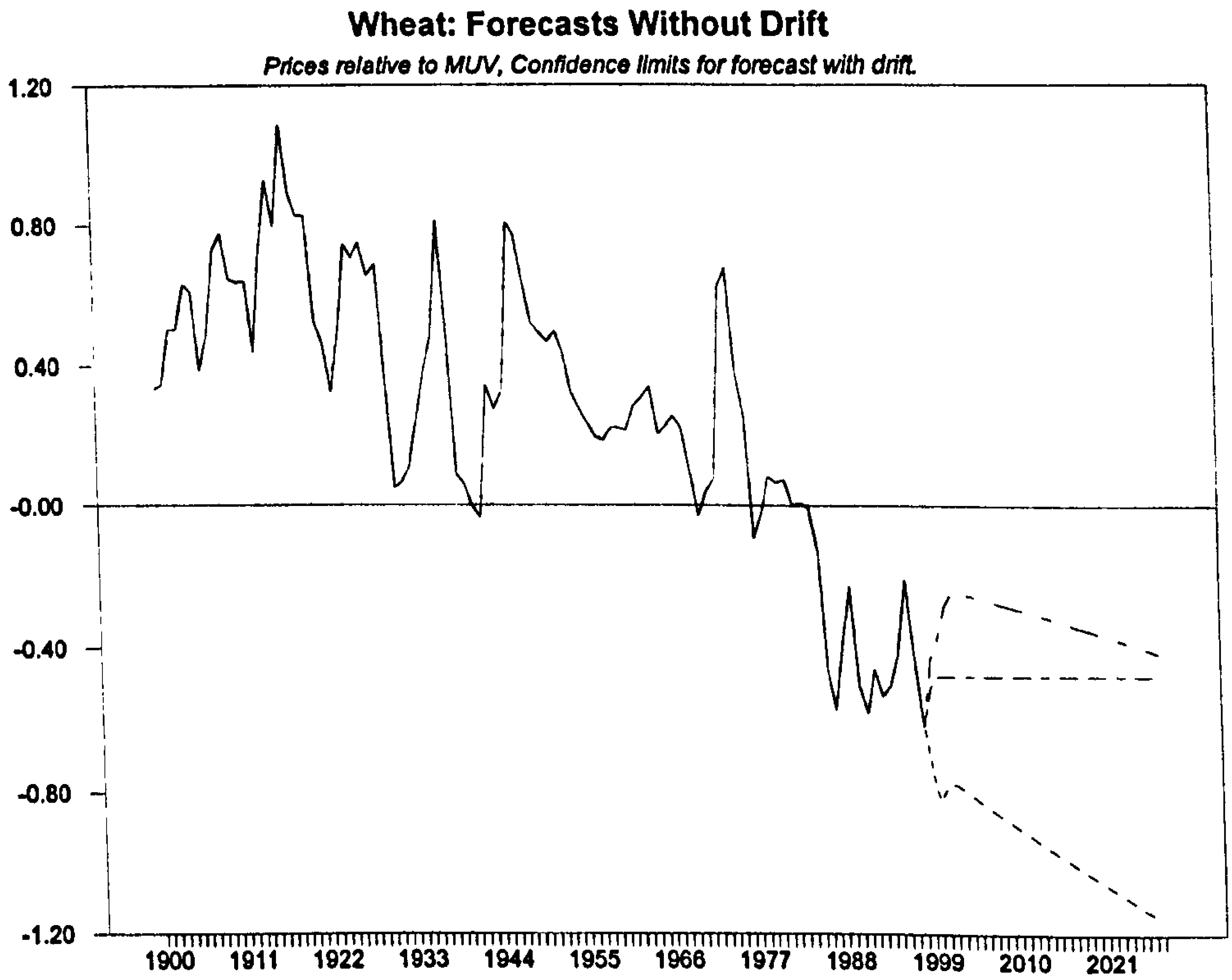
These forecasts and confidence intervals are illustrated for a 20 year forecast horizon in Figure V.ii.i. below.

Figure V.ii.i:



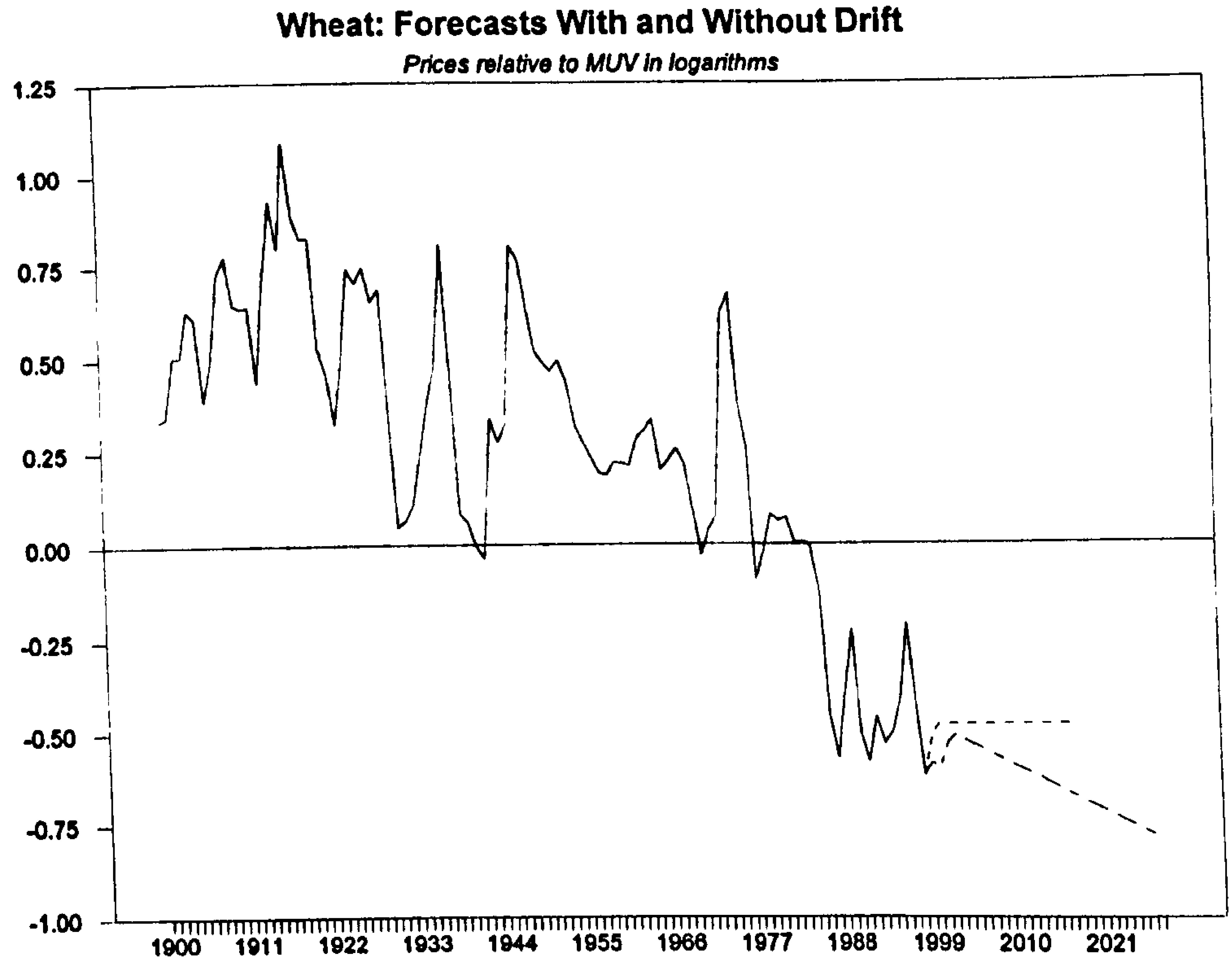
Over the 20 Year forecast horizon shown in Figure V.ii.i, the point forecast from the model without a constant selected by SBC stay within the 68% confidence interval for the forecasts from the model with drift selected by AIC. However the point forecasts from the model without drift come close to the upper limit of the confidence interval of the counterfactual model towards the end of a 30 year horizon. This is illustrated in Figure V.ii.ii below.

Figure V.ii.ii:



Moreover, the point forecasts obtained from both models deviate consistently, as is to be expected. Figure V.ii.iii shows the alternative point forecasts from forecast models with and without drift.

Figure V.ii.iii:



Clearly, the difference between forecast values, while moderate initially, becomes larger as the trajectory of the point forecasts deviates over longer forecast horizons. (The forecast with trend predicts an increase in relative prices to -0.544 over the first seven years[2], and a further decline thereafter. The initial relative price increase appears as a temporary approximation to the forecast from the model excluding drift, both forecasts then deviate from 2005 onwards.) Since, the noticeable increase in the discrepancy between forecast values occurs relatively early the question of selecting between models with and without trend and the persistent uncertainty surrounding this question remains an issue.

[2] The forecasts fluctuate somewhat in either direction, however, it is predicted that a value of -0.544 will be reached after seven years, starting from a final observation of -0.618 and that forecasts will reflect the downwards drift thereafter.

Appendix V.iii. Estimation Results for Relative Primary Product Price Series in First Differences -Minimum SBC Specifications

This appendix gives details of the estimation results for those forecast models where the trend or constant had been dropped and where re-estimation has therefore become necessary without the trend or drift term. Except for Zinc, the models in question are difference stationary models as selected by minimum SBC, but without the constant. Were random walks were selected by SBC, they have been replaced with ARIMA(1,1,1) models here, irrespective of whether an ARIMA(1,1,1) model or a random walk was selected for forecasting (for obvious reasons, since for a pure random walk there would simply be nothing to estimate and hence no estimate to report). The difference stationary model adopted is now:

$$\Delta p_t = v_t$$

and

$$v_t - \phi_1 v_{t-1} - \ldots - \phi_p v_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

where Δp_t is the relevant price series in first differences, ϕ_p the coefficient on the p^{TH} autoregressive error term (v_{t-i}), θ_q the coefficient on the i^{th} moving average term (ε_{t-i}). The subscript t indicates time period t and t-i the i^{th} lag. The general form of the trend stationary model is correspondingly:

$$p_t = a + u_t,$$

with:

$$u_t - \phi_1 u_{t-1} - \ldots - \phi_p u_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

where a is the constant, p_t the relative price level and ϕ_p and θ_q are the autoregressive and moving average coefficients as above.

The estimation results were obtained using the ARIMA.SRC procedure in GAUSS, and are listed below. As in earlier appendices, the Ljung Box Q statistic is reported for 12 autocorrelations, with P-values given in parentheses.

Commodity: Coffee
Model: ARIMA(1,1,1) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**8.516 (0.579)

$$\Delta p_t = v_t$$
$$v_t \underset{(0.072)}{-0.839} v_{t-1} = \underset{(0.243)}{\varepsilon_t} \underset{(0.077)}{-0.993} \varepsilon_{t-1}$$

Commodity: Cocoa
Model: ARIMA(2,1,0) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**5.859 (0.827)

$$\Delta p_t = v_t$$
$$v_t \underset{(0.097)}{-0.083} v_{t-1} \underset{(0.097)}{+0.309} v_{t-2} = \underset{(0.243)}{\varepsilon_t}$$

Commodity: Tea
Model: ARIMA(1,1,1) without constant
Degrees of freedom: 95 , **Ljung-Box Q(12):**13.972 (0.174)

$$\Delta p_t = v_t$$
$$v_t \underset{(0.337)}{+0.692} v_{t-1} = \underset{(0.163)}{\varepsilon_t} \underset{(0.284)}{+0.797} \varepsilon_{t-1}$$

Commodity: Maize
Model: ARIMA(0,1,2) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**12.588 (0.248)

$$\Delta p_t = v_t$$
$$v_t = \underset{(0.205)}{\varepsilon_t} \underset{(0.094)}{-0.192} \varepsilon_{t-1} \underset{(0.094)}{-0.415} \varepsilon_{t-2}$$

Commodity: Beef
Model: ARIMA(1,1,1) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**10.700 (0.381)

$$\Delta p_t = v_t$$
$$v_t \underset{(1.231)}{+0.341} = \underset{(0.207)}{\varepsilon_t} \underset{(1.195)}{+0.408}\varepsilon_{t-1}$$

Commodity: Bananas
Model: ARIMA(1,1,1) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**8.950 (0.537)

$$\Delta p_t = v_t$$
$$v_t \underset{(0.658)}{+0.430}v_{t-1} = \underset{(0.091)}{\varepsilon_t} \underset{(0.614)}{+0.538}\varepsilon_{t-1}$$

Commodity: Palm Oil
Model: ARIMA(2,1,0) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**7.603 (0.668)

$$\Delta p_t = v_t$$
$$v_t \underset{(0.096)}{-0.053}v_{t-1} \underset{(0.096)}{+0.358}v_{t-2} = \underset{(0.213)}{\varepsilon_t}$$

Commodity: Cotton
Model: ARIMA(2,1,2) without constant
Degrees of freedom: 94 , **Ljung-Box Q(12):**9.089 (0.335)

$$\Delta p_t = v_t$$
$$v_t \underset{(0.076)}{-1.306}v_{t-1} \underset{(0.076)}{+0.757}v_{t-2} = \underset{(0.149)}{\varepsilon_t} \underset{(0.062)}{-1.650}\varepsilon_{t-1} \underset{(0.066)}{+0.966}\varepsilon_{t-2}$$

Commodity: Jute
Model: ARIMA(0,1,2) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**13.326 (0.206)

$$\Delta p_t = v_t$$
$$v_t = \underset{(0.214)}{\varepsilon_t} \underset{(0.094)}{-0.046}\varepsilon_{t-1} \underset{(0.095)}{-0.392}\varepsilon_{t-2}$$

Commodity: Wool
Model: ARIMA(0,1,2) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**4.469 (0.924)

$$\Delta p_t = v_t$$
$$v_t = \underset{(0.188)}{\varepsilon_t} \underset{(0.096)}{-0.135}\varepsilon_{t-1} \underset{(0.096)}{-0.381}\varepsilon_{t-2}$$

Commodity: Tobacco
Model: ARIMA(1,1,1) without constant
Degrees of freedom: 95 , **Ljung-Box Q(12):**10.888 (0.366)

$$\Delta p_t = v_t$$
$$v_t \underset{(1.808)}{-0.163} v_{t-1} = \underset{(0.142)}{\varepsilon_t} \underset{(1.823)}{-0.105} \varepsilon_{t-1}$$

Commodity: Copper
Model: ARIMA(1,1,1) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**6.747 (0.749)

$$\Delta p_t = v_t$$
$$v_t \underset{(0.733)}{+0.399} v_{t-1} = \underset{(0.189)}{\varepsilon_t} \underset{(0.693)}{+0.504} \varepsilon_{t-1}$$

Commodity: Tin
Model: ARIMA(1,1,1) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**4.994 (0.892)

$$\Delta p_t = v_t$$
$$v_t \underset{(1.557)}{+0.079} v_{t-1} = \underset{(0.191)}{\varepsilon_t} \underset{(1.545)}{+0.143} \varepsilon_{t-1}$$

Commodity: Silver
Model: ARIMA(2,1,0) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**8.087 (0.620)

$$\Delta p_t = v_t$$
$$v_t \underset{(0.097)}{-0.042} v_{t-1} \underset{(0.097)}{+0.307} v_{t-2} = \underset{(0.180)}{\varepsilon_t}$$

Commodity: Lead
Model: ARIMA(1,1,1) without constant
Degrees of freedom: 96 , **Ljung-Box Q(12):**9.735 (0.464)

$$\Delta p_t = v_t$$
$$v_t \underset{(0.162)}{-0.726} v_{t-1} = \underset{(0.183)}{\varepsilon_t} \underset{(0.113)}{-0.900} \varepsilon_{t-1}$$

Commodity: Zinc
Model: ARIMA(1,0,1) with constant and trend
Degrees of freedom: 96 , **Ljung-Box Q(12):**6.390 (0.781)

$$p_t = -\underset{(0.047)}{0.020} + u_t$$
$$u_t - \underset{(0.132)}{0.430} = \underset{(0.195)}{\varepsilon_t} + \underset{(0.138)}{0.380}\varepsilon_{t-1}$$

Chapter 6

Trend-Cycle Decompositions for Individual Commodity Price Series

Chapter 6: Trend-Cycle Decompositions for Individual Commodity Price Series

Previous chapters have focused on the presence, the magnitude and the sign of trend components in relative commodity price series as a basis for formulating forecast models. It has been obvious throughout, though, that irrespective of the presence of a trend or drift most of these relative price series are characterised by substantial volatility. This chapter will provide illustrations of the trend and cycle components of the relative commodity price series considered.

6.1. Trends and volatility in commodity price series

The incidence and pattern of commodity price volatility has in itself been the subject of extensive research. The question of volatility was considered alongside the issue of long run trends in Borzenstein *et. al.* (1994), Reinhart and Wickham (1994) and León and Soto (1997). A number of recent IMF studies -such as Cashin *et. al.* (1999b), Cashin and Patillo (2000) and more recently: Cashin and McDermott (2001)- have considered the question of commodity price fluctuations with respect to magnitude, duration and symmetry characteristics.

While this separate field of research is too extensive to be treated here, it is worth remembering that among the studies quoted above Cashin and McDermott (*op. cit.*), using the Economist Industrial Price Index, confirm that commodity prices have fallen over time, although they also argue that the development of the price series is more appropriately characterised by large fluctuations.

A number of studies looking into the significance and magnitude of trend components in commodity price series have also covered trend cycle decompositions of the price series used. Cuddington and Urzúa (1989) compute the Beveridge Nelson Decomposition for the aggregate Grilli and Yang index, while Borzenstein *et.al.* (1994), using quarterly commodity price data, supply Beveridge Nelson decomposition results for sub indices covering Food, Metals and Beverages as well as a decomposition for a composite index of all commodity price series considered by the authors. Cuddington (1992) finally computed Beveridge Nelson decompositions for those of the individual Grilli and Yang commodity price series for which a difference stationary specification was chosen through *a priori* testing for unit roots.

Borzenstein *et.al.* report relatively low shock persistence for metals and high shock persistence for beverages, with food prices taking an intermediate position. However, no such general observation can be made on the basis of the results by Cuddington (1992): Cuddington (*op. cit.*) reports persistence results for a number of price series modelled as difference stationary. Although he reports high shock persistence for difference stationary metal price series (Aluminium 88.3%, Copper 100%, Silver 64.1%) similarly high persistence results are recorded for a number of food commodities (Beef 100%, Cocoa 64.1% and Tea 72%). Only the shock persistence for Banana prices is inferred to be markedly lower with 44.4%. Among other commodities shock persistence is high for Rubber (100%), Tobacco (78.1%) and arguably Cotton (56.1%). Noticeably lower shock persistence is inferred for Jute (40%) and Wool (34.3%).

In the present study, Beveridge Nelson Decomposition results are presented for all those commodities where a difference stationary model can plausibly be considered[1]. This will be preceded by a brief review of the underlying methodology and the computation method employed.

6.2. Fundamentals of the Beveridge Nelson Decomposition Method

Following Beveridge and Nelson (1981), an integrated time series can be decomposed into a permanent and cyclical component, allowing for a constant and general ARMA(p,q) errors in the underlying data generating process. In their original article (Beveridge and Nelson (*op. cit.*)) the authors represent their decomposition result in terms of a pure moving average process. This approach clearly would encompass the case of inverted autoregressive components. However, a formulation involving only a finite number of parameter estimates is clearly desirable for the purposes of practical implementation. Furthermore, autoregressive components in the residual process of the ARIMA models considered have been explicitly modelled as such in previous parts of this study. An exact computation method for the Beveridge Nelson decomposition on the basis of an ARMA(p,q) representation of the differenced series has been introduced by Newbold (1990). It appears desirable with a view not only to accurate implementation but also in the interest of consistent presentation, to follow the treatment in Newbold (1990) throughout. Thus, for a price series p_t:

[6.2.1] $$w_t = p_t - p_{t-1}$$

[1] In the case of some commodity price series where support for a trend stationary or stationary representation is very strong the results of a Beveridge Nelson decomposition would merely confirm the inadequacy of the difference stationary model. In those cases decomposition or shock persistence results will only be mentioned if it is informative to do so for comparative purposes.

with residual $v_t = w_t - \beta$, where β is as before the constant representing the drift term, one obtains[2]:

[6.2.2] $$\phi(L)(w_t - \beta) = \theta(L)\varepsilon_t$$

It then follows immediately from [6.2.1] that $p_t = w_t + p_{t-1}$ or, more generally, that the value of p in a given time period $t+k$ can be obtained from a past value of p_t and the sum of the intermittent one period changes in the series, so that forecasting from t one has:

[6.2.3] $$\widehat{p}_t(k) = p_t + \sum_{j=1}^{k} \widehat{w}_t(j)$$

where $\widehat{p}_t(k)$ is the k period ahead forecast of p_t, and $\widehat{w}_t(j)$ the projected j period ahead change in p_t.

The differenced series w_t is stationary so long as p_t is I(1) and is asymptotically linear in k, the length of the forecast horizon. The slope of this function is given by the drift parameter β and its intercept is given by:

[6.2.4] $$\bar{p}_t = p_t + \lim_{k\to\infty}\left[\sum_{j=1}^{k} \widehat{w}_t(j) - k\beta\right]$$

Defining the cyclical component as:

[6.2.5] $$c_t = \bar{p}_t - p_t$$

The original price series is seen -under the Beveridge Nelson definition of the cycle- as the difference between the permanent and the cyclical component (*i.e.* $p_t = \bar{p}_t - c_t$). As pointed out in Newbold (1991), an intuitively more appealing representation follows if one re-expresses the cycle as $I_t = -c_t$ and hence the data series as the sum of trend and cycle ($p_t = \bar{p}_t + I_t$).

[2] The exposition of the Beveridge Nelson Decomposition method presented here follows the summary given in Newbold (1990), although the denomination of the variables involved has been amended to maintain consistency with the notation in previous chapters.

Two points are worth noting about the Beveridge Nelson decomposition technique: First, the cyclical component (c_t) is defined exclusively in terms of the ARMA(p,q) parameterization of the residual process and does not refer to any kind of regularity which would be expected to have a counterpart in systemic forces of adjustment endogenous to a given economic model (*cf.* Beveridge Nelson (*op. cit.*) and Newbold (1991) where the adequacy of this definition is also discussed further with a view to its statistical properties).

A second point of interest are the characteristics of the *permanent* component. In spite of its name, it is pointed out by Beveridge and Nelson (*op. cit.*) that the extracted permanent component does itself follow a random walk, so that the permanent value can differ substantially at different values of t. (Of course it also follows that the decomposition method can not be applied to a difference stationary series that follows a random walk to start with. By definition the Beverdige Nelson decomposition would merely return the original data series in such a case.) As a counterpart of this characteristic, the cyclical series c_t is stationary by construction. This should be obvious from the fact that the Beveridge Nelson definition of the cycle excludes the drift term as well as the permanent effect of innovations on the original data series.

The computational method used is the one derived in Newbold (1990), where it is shown that the exact value of the cyclical component can be obtained from:

[6.2.6] $$c_t = \sum_{j=1}^{q} \hat{y}_t(j) + \frac{\sum_{j=1}^{p} \sum_{i=j}^{p} \phi_i \hat{y}_t(q-j+1)}{1 - \sum_{i=1}^{p} \phi_i},$$

where $y_t = w_t - \beta$ and $\hat{y}_t(j)$ is the corresponding j-step ahead forecast. p and q in [6.2.6] refer to the p autoregressive and q moving average parameters respectively. In implementing [6.2.6], the ϕ_i have been replaced by the estimated coefficients $\hat{\phi}_i$ from the appropriate ARIMA(p,1,q) models discussed previously.

Given the decomposition method's reliance on lagged residual values, some initial observations will invariably be lost in the extracted permanent and cyclical components. In the present case, the initial five observations have therefore uniformly been dropped for all commodity price series in question[3] for the Beveridge Nelson Decomposition results.

In addition to this decomposition, shock persistence in an integrated series can be represented in a steady state gain function, as discussed in Cuddington and Winters (1987), where appropriate[4]. Following Cuddington and Winters (*op.cit.*), the Steady State Gain Function for each of the difference stationary price series is given by:

[6.2.7.] $$\Delta\bar{p}_t = \beta + \frac{(1-\theta_1-\theta_2-\cdots-\theta_q)}{(1-\phi_1-\phi_2-\cdots-\phi_p)}\varepsilon_t$$

where ϕ_i and θ_j with $i = 1,2,\cdots,p$ and $j = 1,2,\cdots,q$ are again replaced with the estimated $\hat{\phi}_i$ and $\hat{\theta}_j$ parameters from the ARIMA models presented in previous chapters.

[3] In some cases, *e.g.* where the selected model was ARIMA(0,1,2) as few as two observation would have been needed to be dropped. However in the interest of consistent presentation as well as to allow some extra observations for models containing autoregressive terms, it was deemed appropriate to present decomposition results for data from 1905 onwards.

[4] Use of the gain function is here deemed appropriate where the Beveridge Nelson decomposition was, *i.e.* in those cases where the use of integrated models either for forecasting purposes or for backwards looking analysis should be considered.

Cuddington and Winters (*op. cit.*) refer to Box and Jenkins (1976)[5] where it is shown that the steady state gain from a pure ARMA or ARIMA process can be represented by a ratio of two polynomials of the lag operator. In the context of [6.2.7] this implies that this ratio of two polynomials defines the steady state gain up to a positive constant, from which it can be seen in turn that:

[6.2.8] $$\zeta = \frac{(1-\theta_1-\theta_2-\cdots-\theta_q)}{(1-\phi_1-\phi_2-\cdots-\phi_p)}$$

describes the proportion of a random shock to an ARIMA model which has a permanent impact on the gain function.

6.3. Trend cycle decomposition results for integrated series

For those models where an ARIMA(p,1,q) representation appears appropriate, illustrations of Beveridge Nelson trend cycle decompositions, .following the methodology outlined above under 6.2. are presented subsequently. Focusing on cereals initially, the permanent components and original data series for Rice, Wheat and Maize, corresponding to the Beveridge Nelson decomposition for their minimum SBC models[6] are illustrated in Figures 6.3.1 to 6.3.3:

The figures reflect the persistence results from the gain function according to [6.2.7] and [6.2.8] well. The highest shock persistence among the three cereals is noted for Wheat where 52.021% of a shock to the series have a permanent effect, while only 34.068% of a shock to the relative price series of Maize are classified as permanent. For Rice the permanent component of a shock is even lower with

[5] Box and Jenkins (1976), pp. 338-346

[6] It will be recalled that these are ARIMA(1,1,2) for Rice and ARIMA(0,1,2) for Wheat and Maize. Since the evidence for the presence of a drift component in the price series for Wheat was detected only after reparameterising the estimated model through selection by AIC, the decomposition corresponding to this model specification is given in Appendix VI.i. for comparative purposes.

24.247%. In figures 6.3.1 to 6.3.3 this is reflected by the fact that the permanent component for Wheat prices is close to the original series, denoting only small cyclical fluctuations while the permanent component for Rice appears more stable and the cyclical deviations are correspondingly large; figure 6.3.3 finally shows the intermediate case for Maize.

Figure 6.3.1

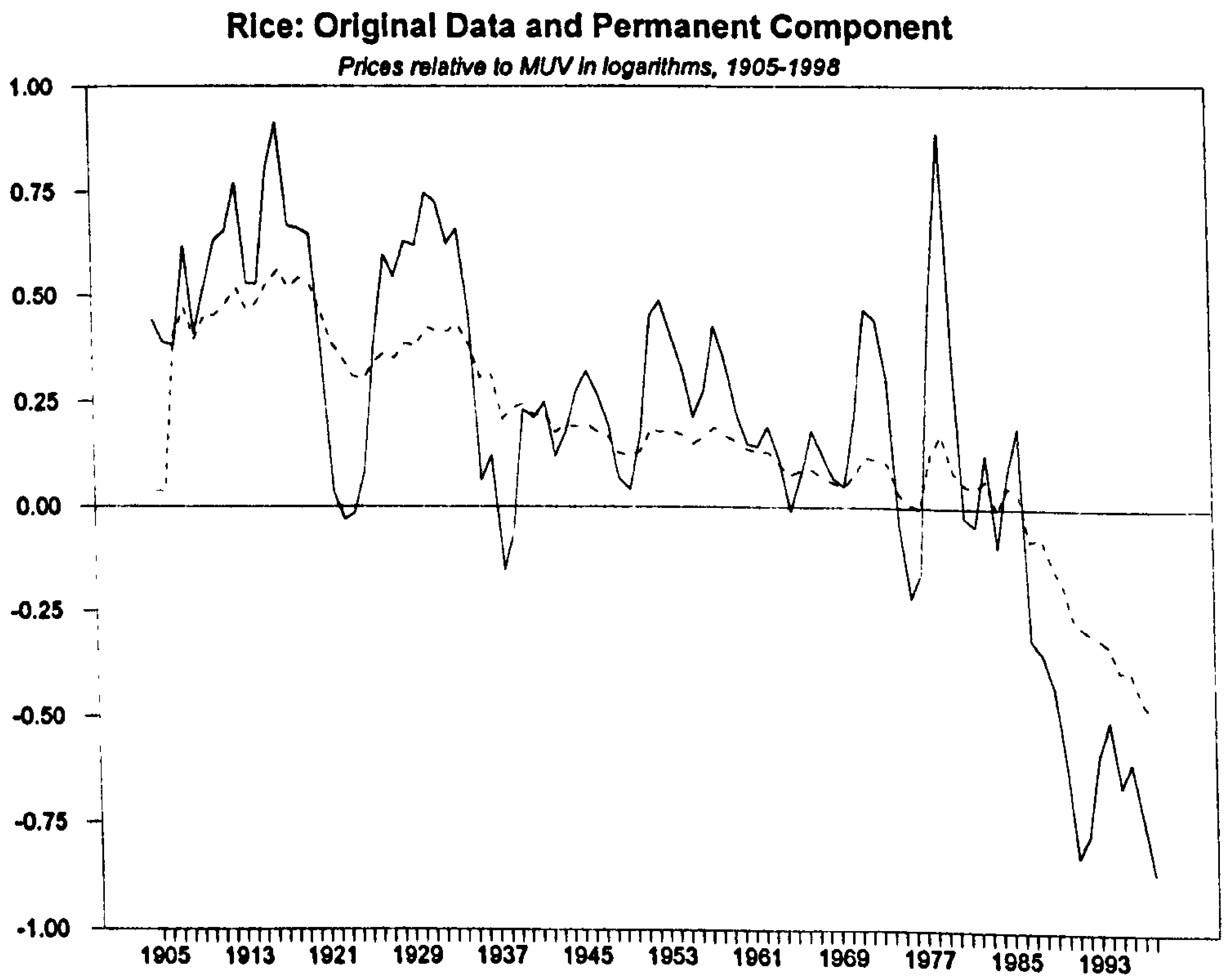


Figure 6.3.2

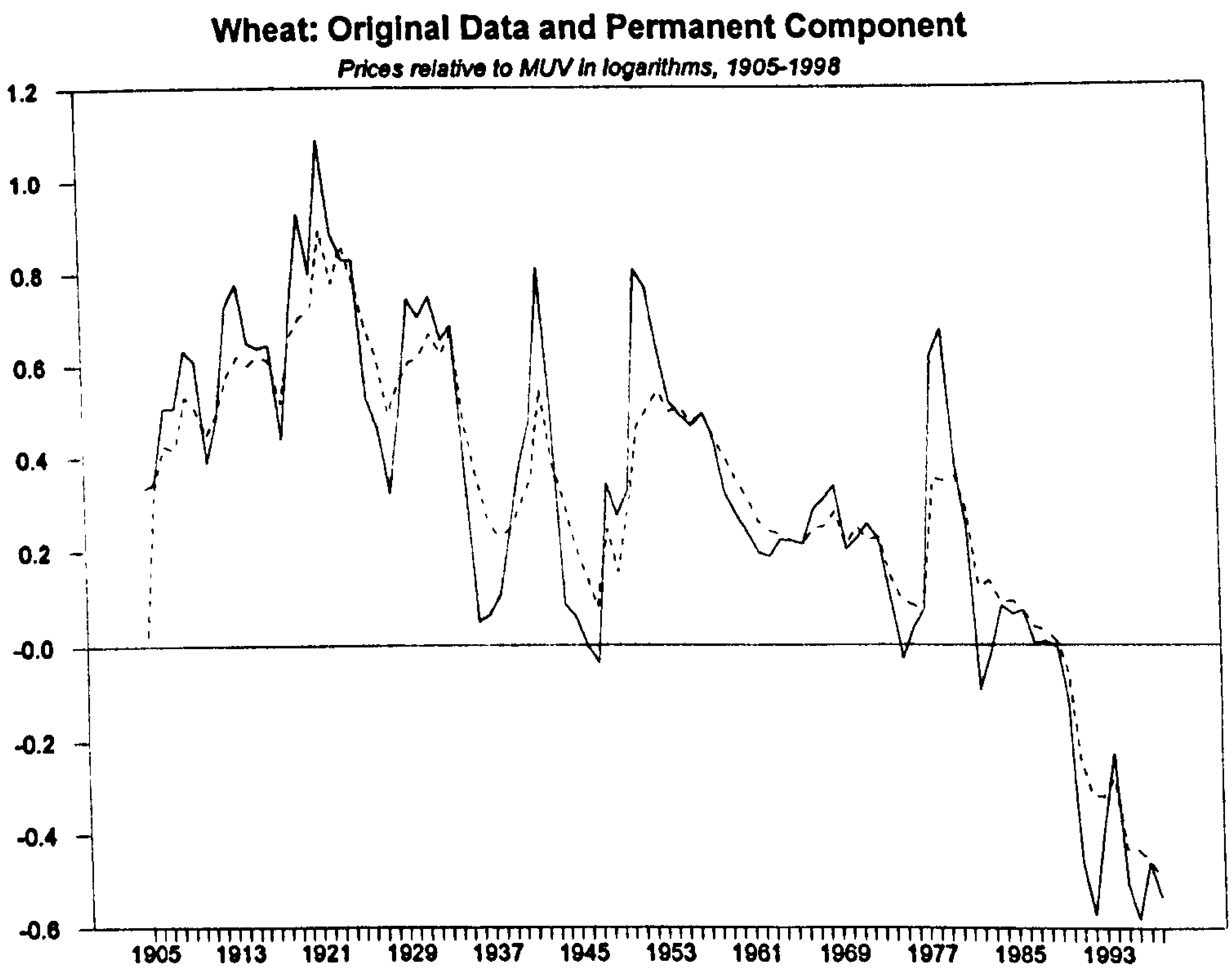
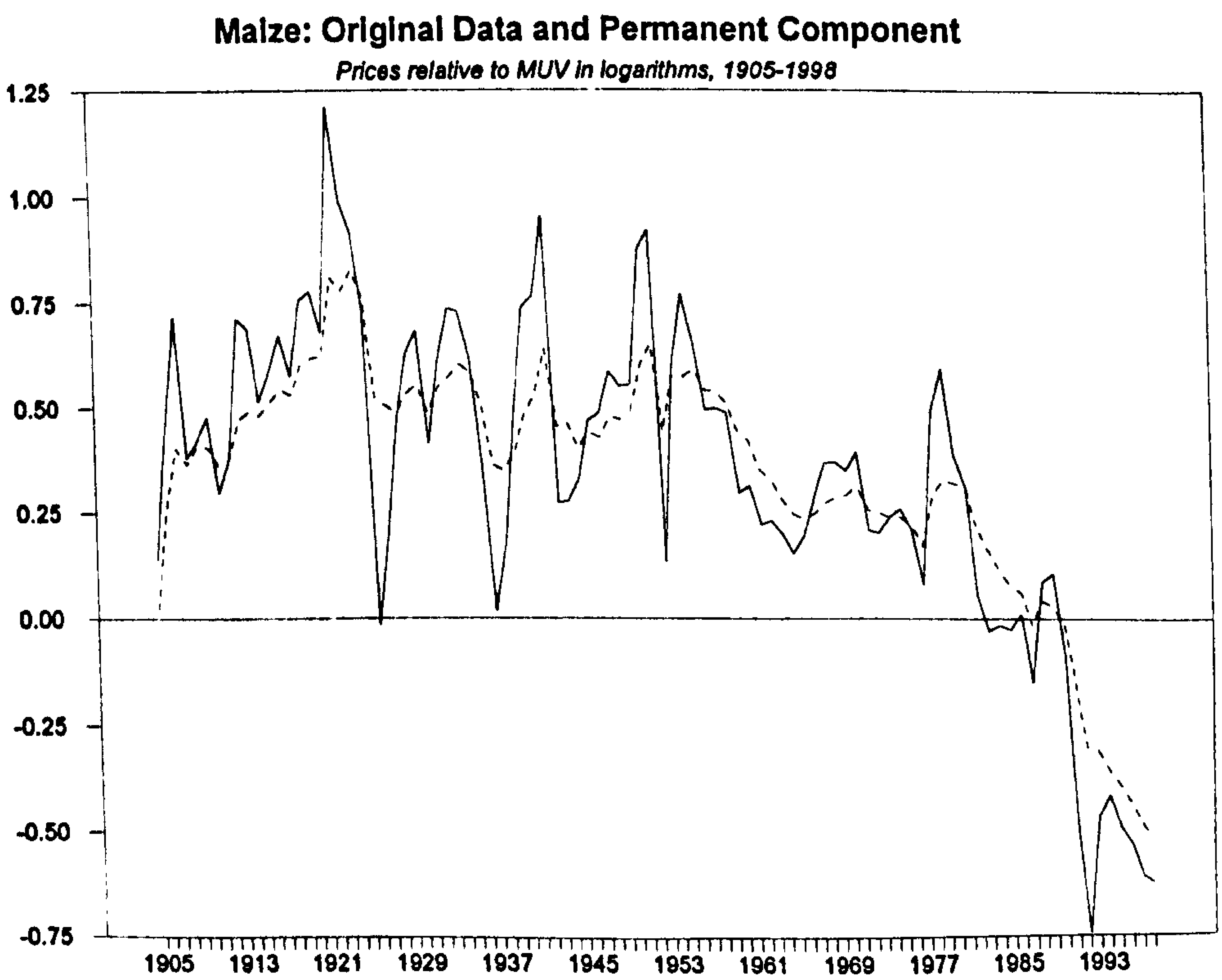


Figure 6.3.3



Other primary commodities such as Cocoa and Palm Oil show high shock persistence[7] and, correspondingly, a permanent component -by the Beveridge Nelson definition- that closely follows the original data series. The permanent component of a random shock to the series is inferred to be 81.365% and 76.457% for Cocoa and Palm Oil respectively. The trajectory of the permanent and transitory components of both series, which are both modelled as ARIMA(2,1,0), are shown in Figures 6.3.4. and 6.3.5. below.

The case is similar for one series of metal prices: in the case of Silver, which is also modelled as ARIMA(2,1,0) the permanent component of a shock to the system is estimated to be 79.010%. The permanent and cyclical components are illustrated in Figure 6.3.6.

[7] While shock persistence can be described as high for one series compared to another, one needs to be aware that shock persistence for a correctly specified model can not be above one or 100%. In this sense higher shock persistence as used in this chapter refers to persistence closer to one.

Figure 6.3.4:

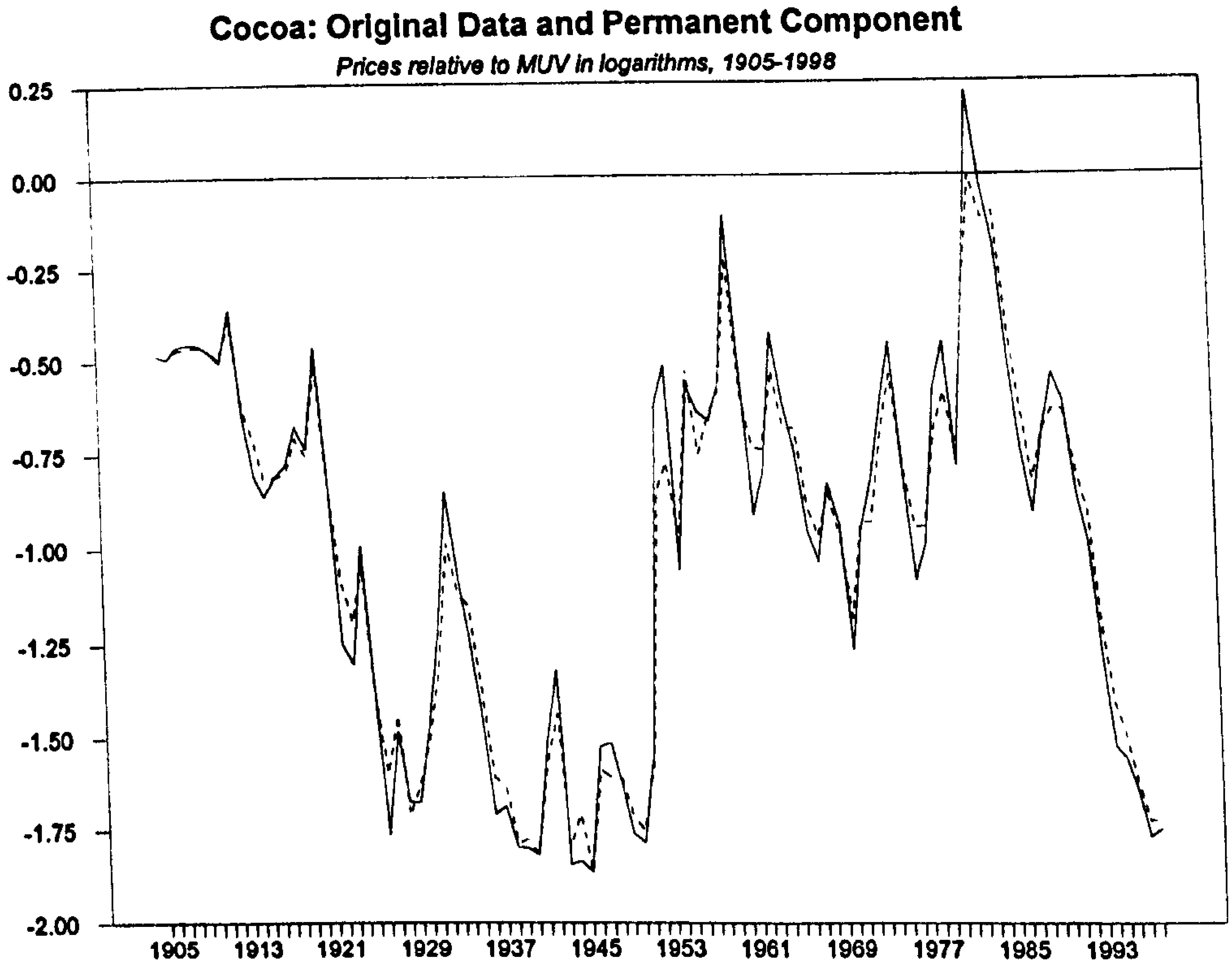


Figure 6.3.5:

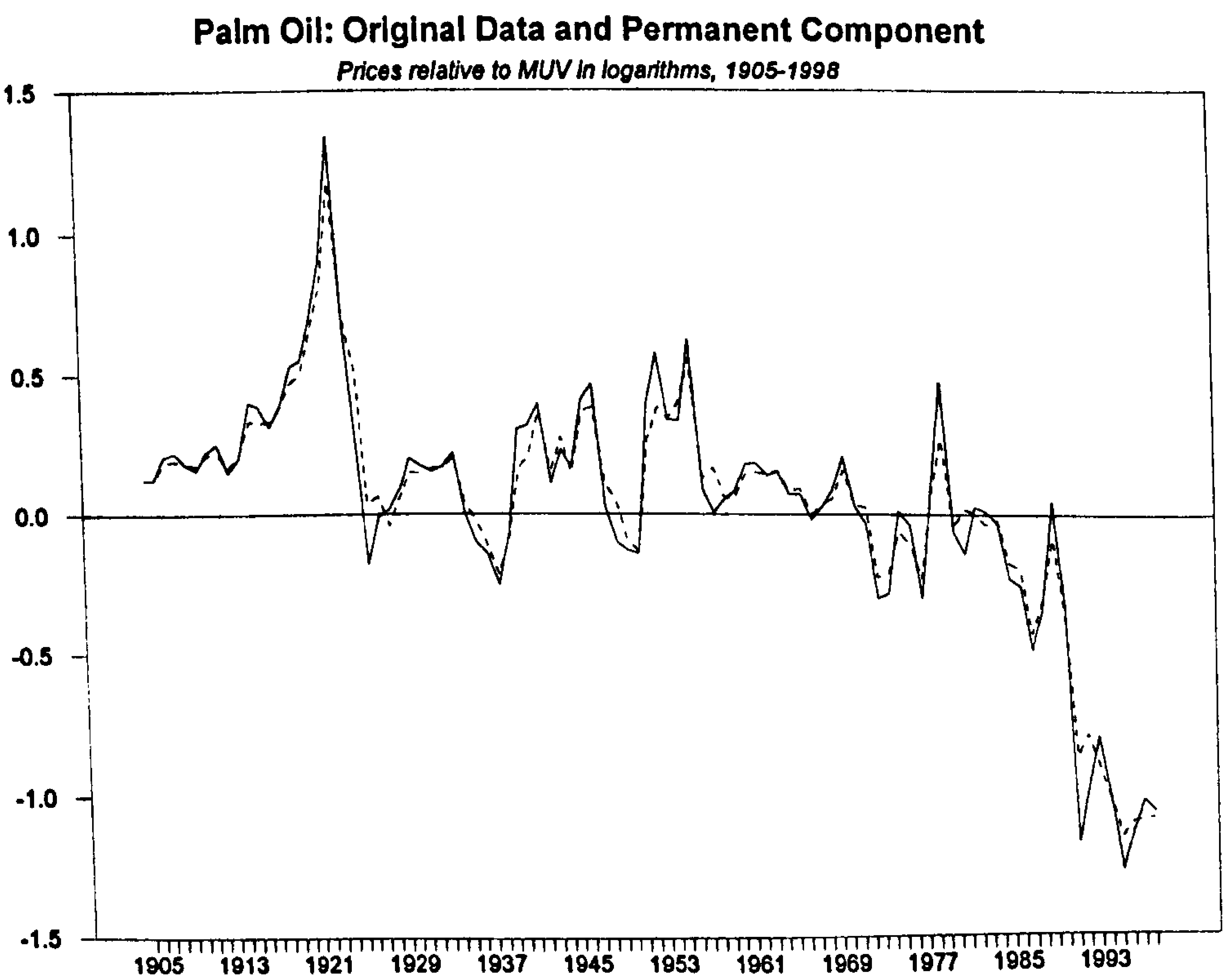
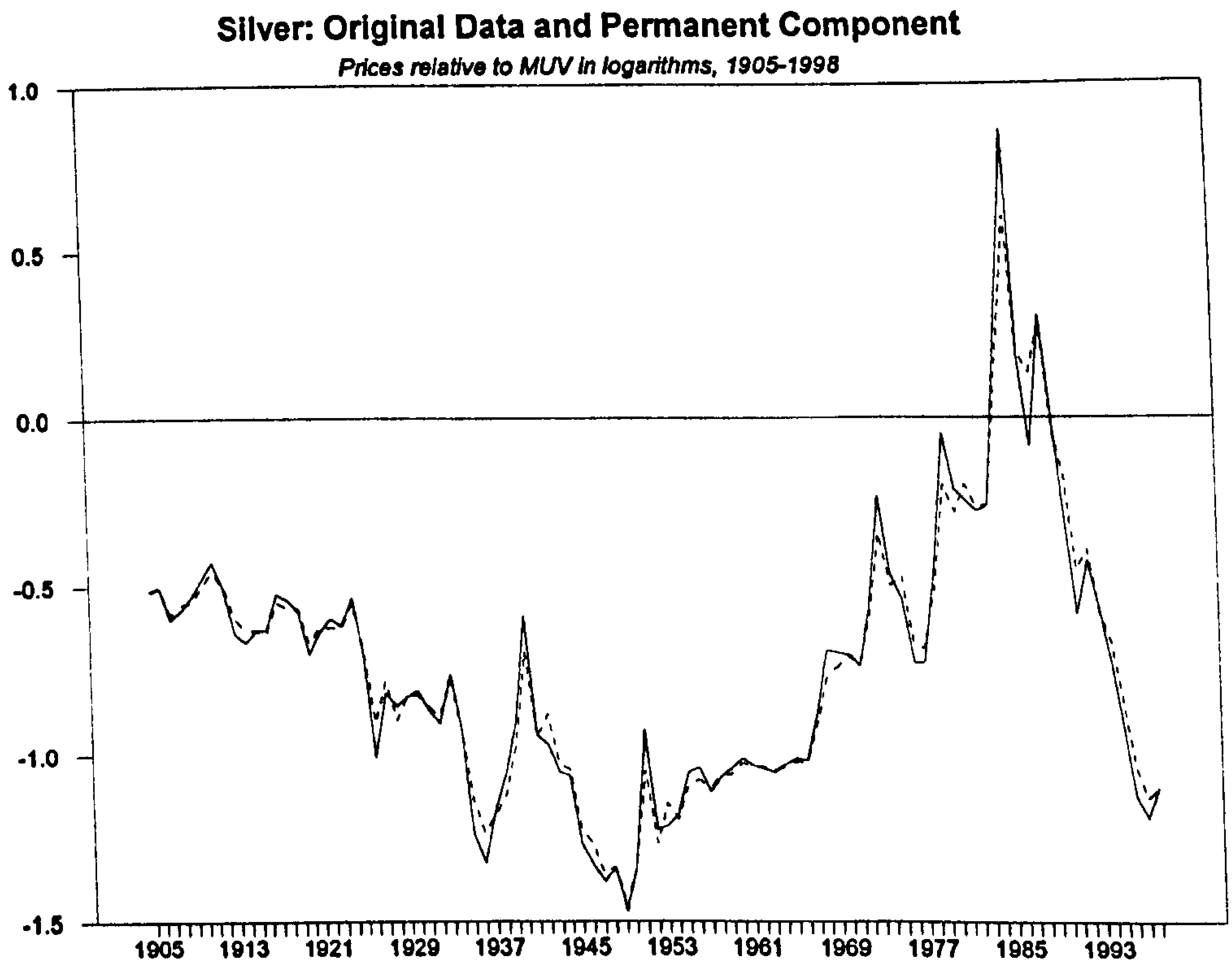


Figure 6.3.6:



Somewhat lower shock persistence is observed for Jute and Wool. 54.781% of a random shock are classified as permanent in the case of Jute and 40.800% in the case of Wool. Shock persistence is higher in the case of Cotton, where 69.411% of a shock to the series are classified as permanent. The permanent component and original data series for Jute and Wool are illustrated in Figures 6.3.7 and 6.3.8 respectively, the higher shock persistence in the case of Jute compared to Wool can be readily appreciated. Figure 6.3.9 shows the permanent component and original data for Cotton, in which case shock persistence is relatively high and the permanent component follows the original data series closely.

Figure 6.3.7:

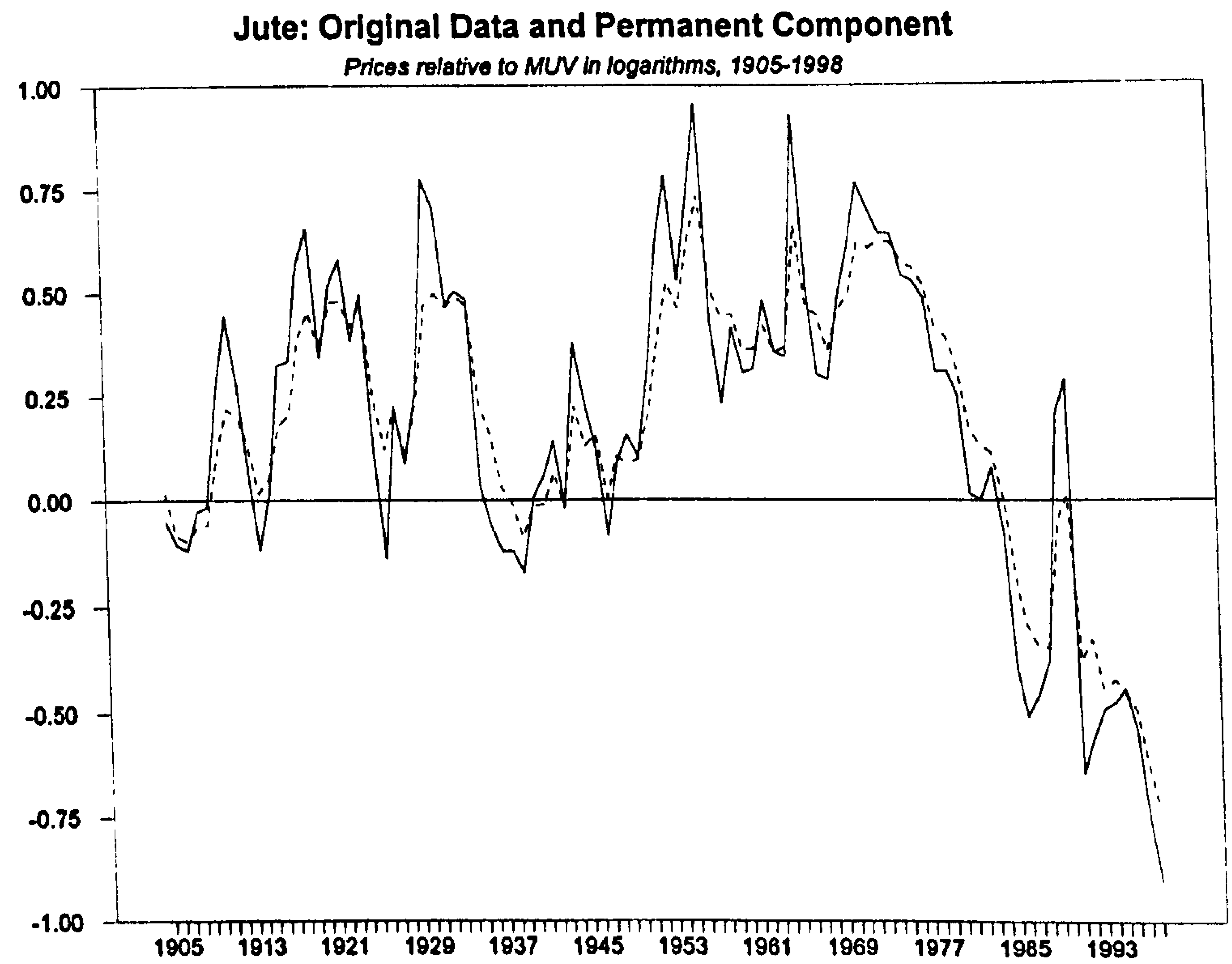


Figure 6.3.8:

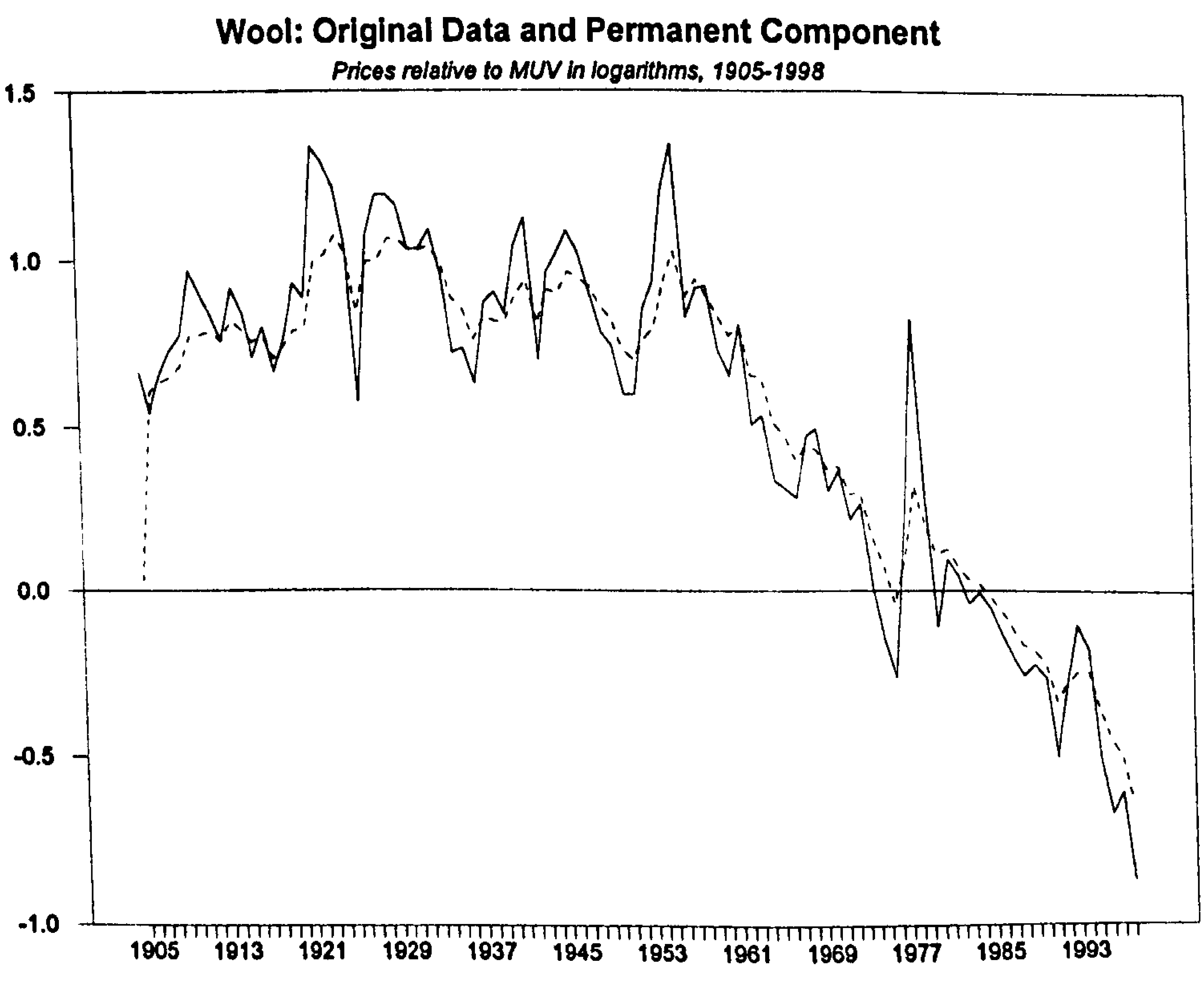
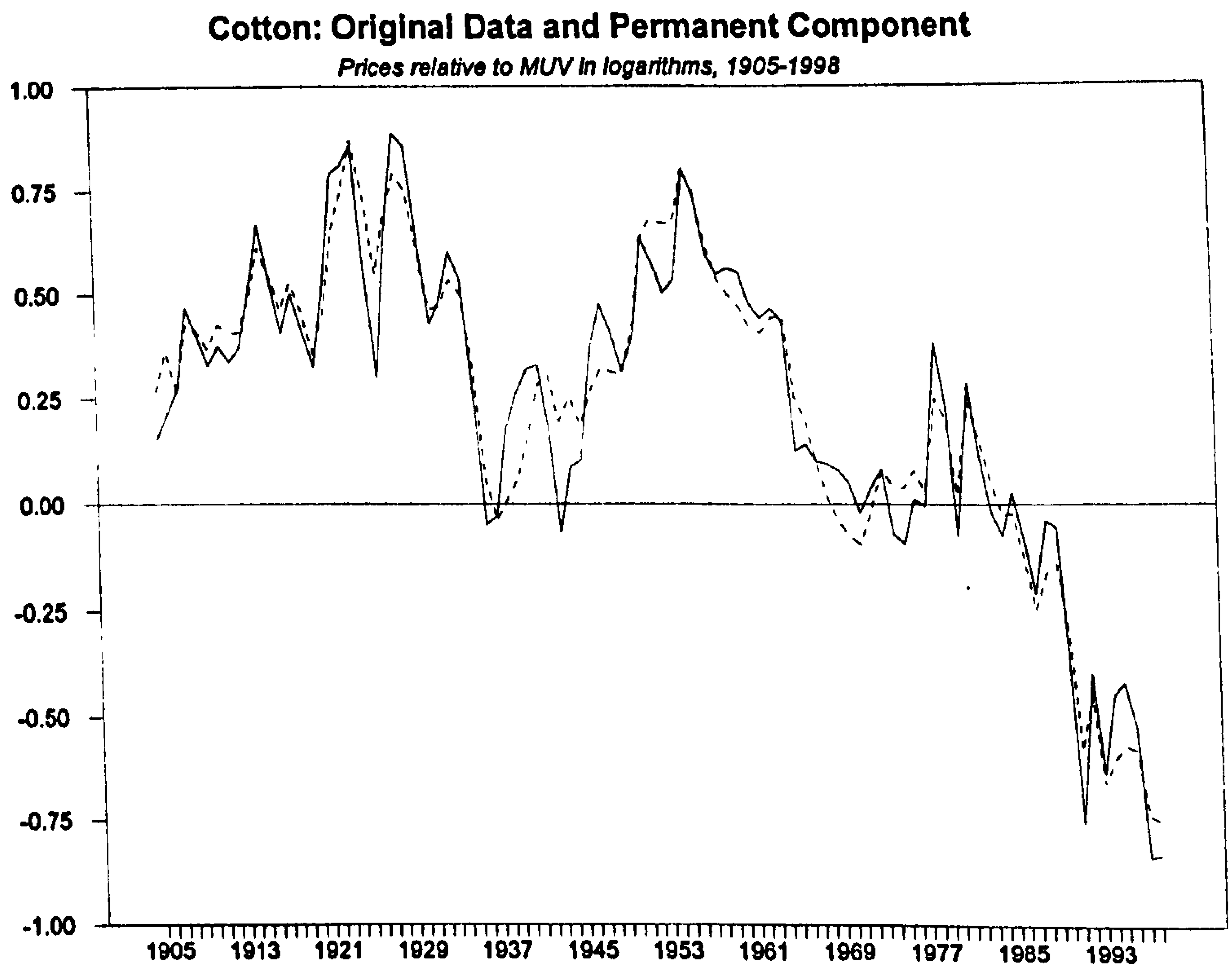


Figure 6.3.9:



Among the remaining price series, the permanent component of a shock to the series was inferred to be 0.000% for Aluminium and Zinc, further confirming the adequacy of the stationary model selected for forecasting purposes.

Random walks had been selected by SBC for the following series: Coffee, Beef, Lamb, Bananas, Rubber, Timber, Copper, Tin, Lead, Tea and Tobacco. Superimposing an ARIMA(1,1,1) on these random walks, the absence of a permanent shock component is established for Coffee, Lamb, Rubber and Timber, while the permanent component of a shock to the series for Lead is estimated to be 0.003%. In all of these five cases the estimated value of the coefficient on the moving average parameter is on the invertibility boundary. Both the low or zero values for the estimated permanent shock component and the point estimates for the MA(1) term support the notion of a stationary series. This is much in accordance with the preceding analysis in the cases of Rubber and Timber. For Lamb, there was substantial uncertainty regarding the order of integration at the point of forecast model selection, although the persistence results and the parameter estimates on the ARIMA(1,1,1) model do support the notion of a stationary model in accordance with the forecast model selection undertaken in Chapter 5. In the cases of Coffee and Lead, the calculations of the gain function for the ARIMA(1,1,1) model also seem to support the notion of a stationary process, although in both cases a difference stationary forecast model had been found to be appropriate in the light of the preceding simulation evidence on the benefits of unit root pre-testing and the likely cost misspecification of the order of integration in forecast models.

For the remaining series (Beef, Bananas, Copper, Tin, Tea and Tobacco) imposing an ARIMA(1,1,1) model would imply a permanent shock component of more than 100% in all cases. Applying the Beveridge Nelson decomposition and plotting the implied permanent component with the original data shows both series to coincide almost exactly. (This is illustrated for the case of Copper in Appendix VI.i.) One can therefore conclude that the series concerned are indeed best represented as a random walk.

6.4. Trend and cycle components in stationary and trend stationary series

For the trend stationary series in the sample (Sugar, Lamb, Rubber, Timber and Aluminium) the estimated trend line is shown together with the original data series below. Figures 6.3.1. to 6.3.3 show the inferred trend lines for Rubber, Timber and Aluminium. It is easily seen that, at least in the latter half of the sample period, the respective price series are well characterised by the fitted trend line, although there is still substantial volatility around it, and deviations from the secular trend have lasted for long periods. Large cyclical deviations over five or more years can be observed for all three commodity series.

Figure 6.4.1:

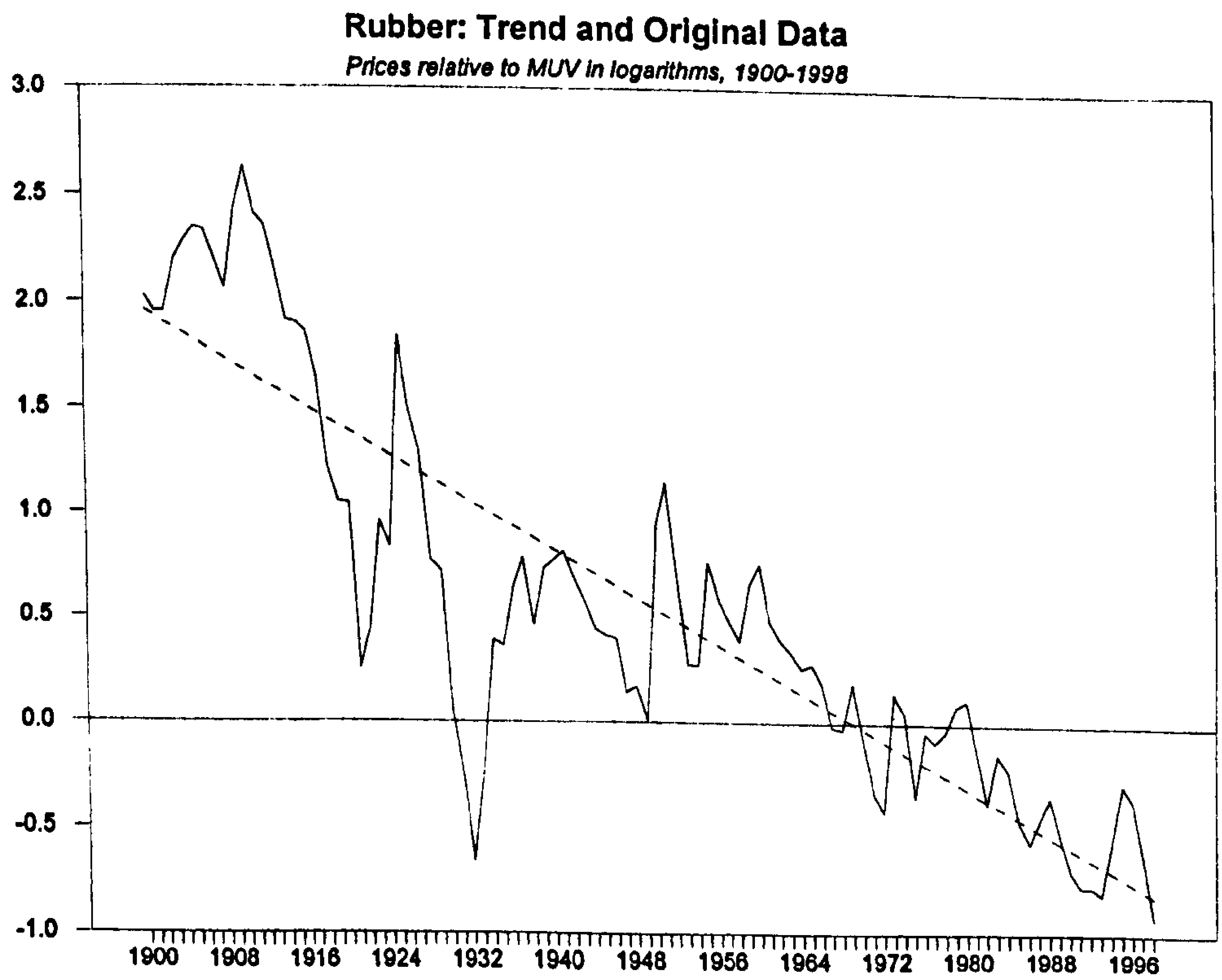


Figure 6.4.2:

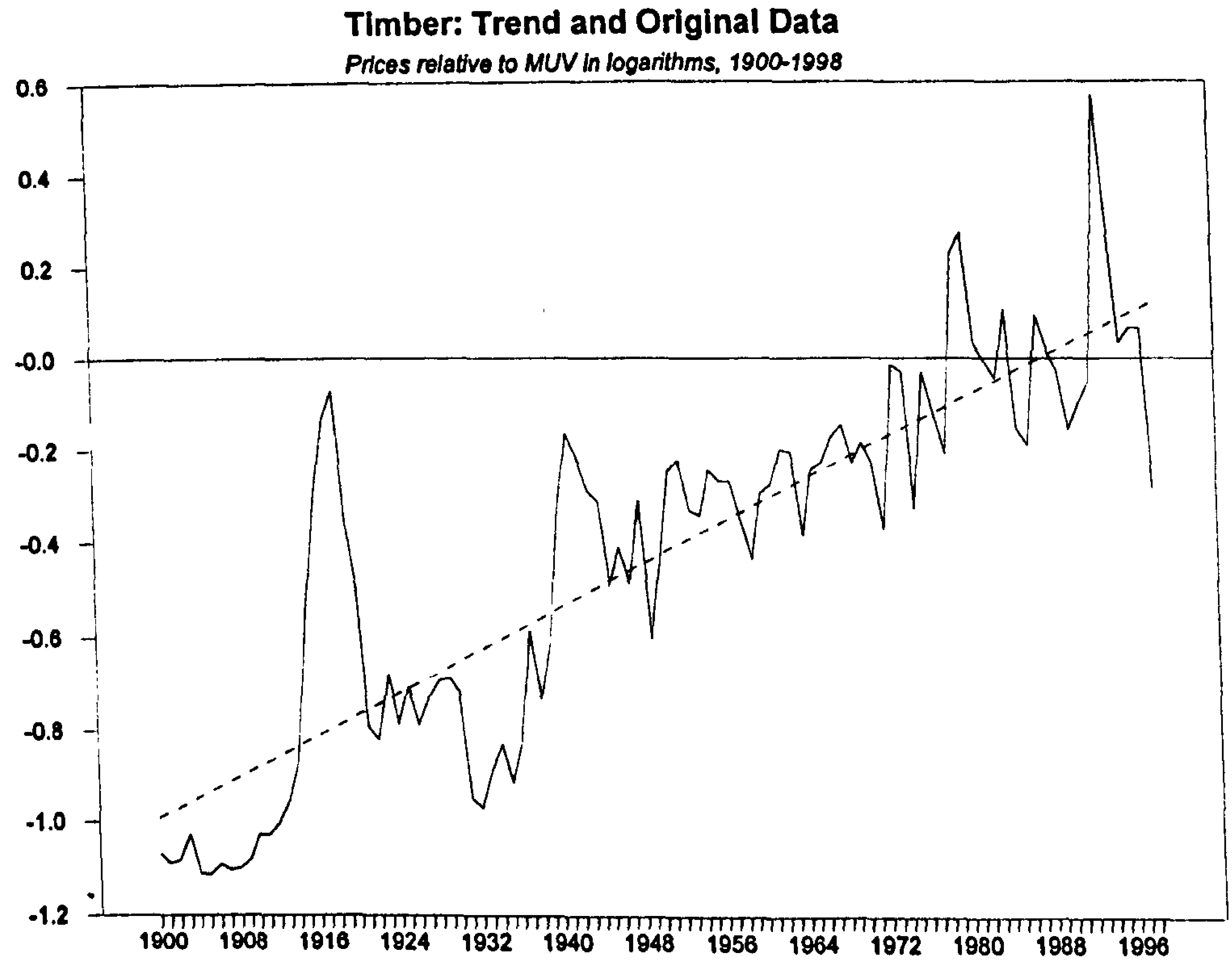
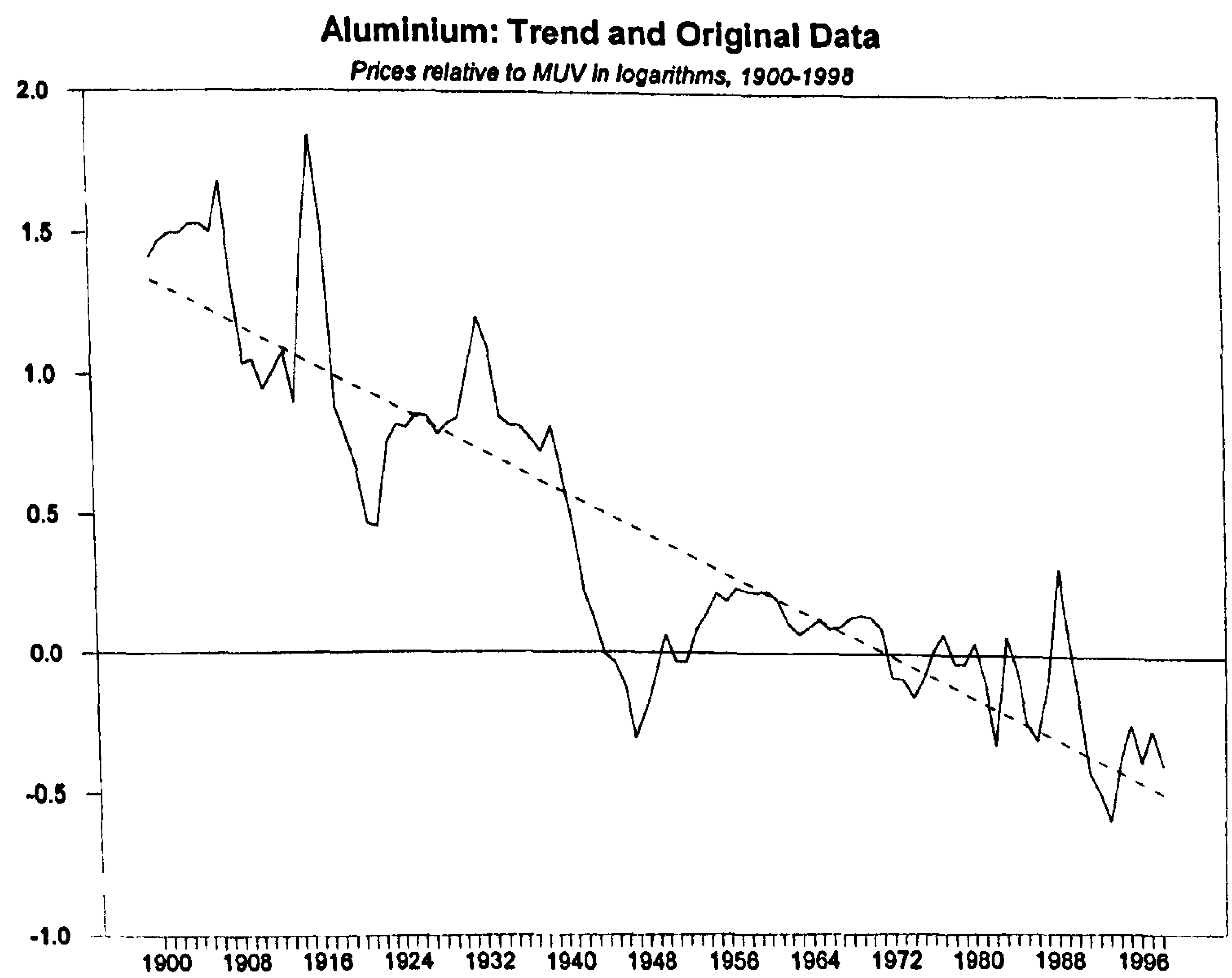


Figure 6.4.3:



This feature is even more pronounced in the cases of Sugar and Lamb, where fluctuations around the secular trend appear to be at least as characteristic a feature of the data as the secular trend line itself. This is illustrated in Figures 6.3.4 and 6.3.5 below.

Figure 6.4.4:

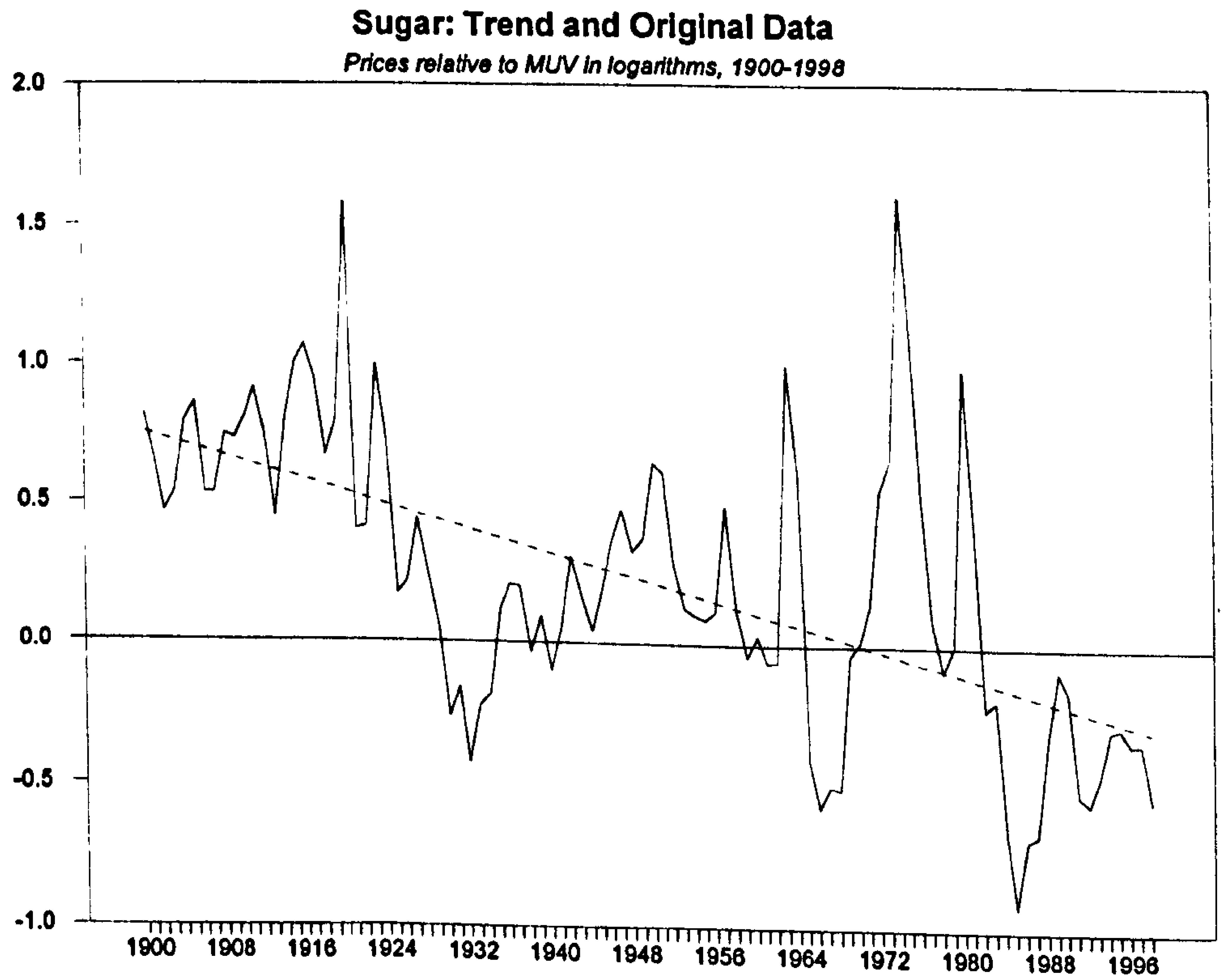
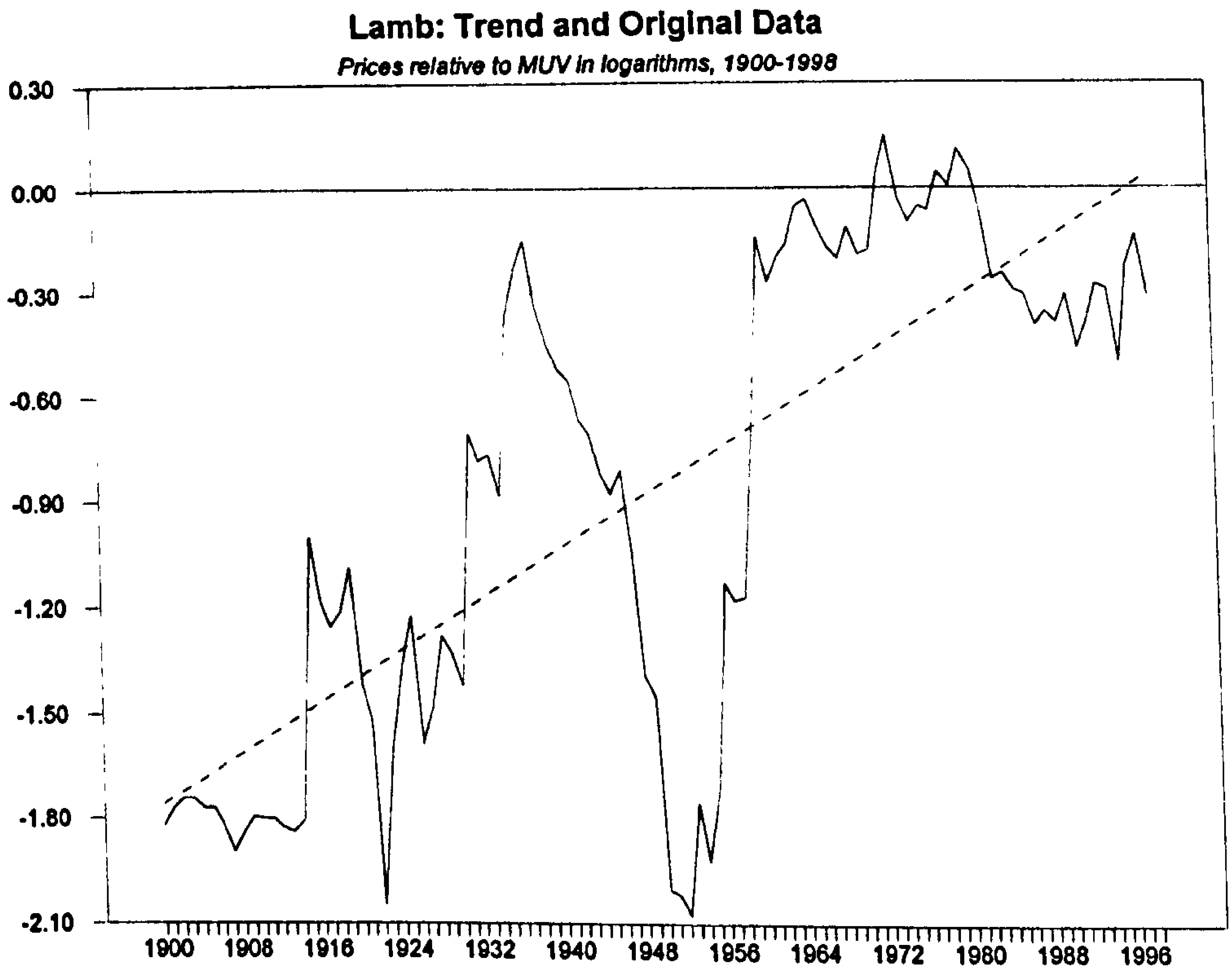


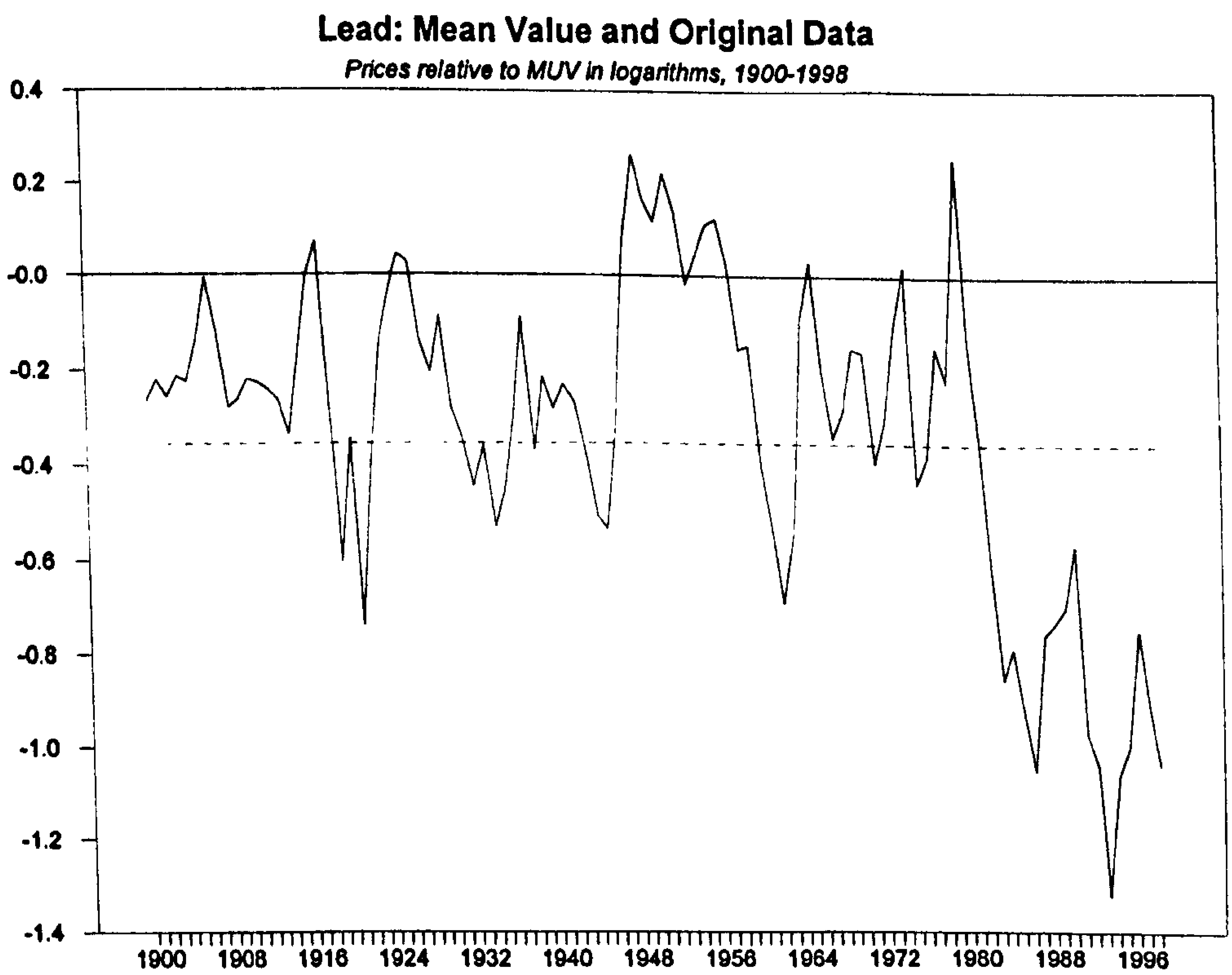
Figure 6.4.5:



The difference in volatility as compared to the trend component is partly reflected in the values of the normalised trend coefficients as reported in Chapters 4 and 5. The normalised trend coefficient for Sugar takes a value of -0.034 compared to 0.076 for Timber, -0.104 for Rubber and -0.124 in the case of Aluminium. In the case of Lamb, the value of the normalised trend coefficient estimate is relatively high with a value of 0.094. Yet a visual inspection of Figure 6.4.5 shows that the presence of large and persistent shifts in the data series make the alternative of an integrated series appear plausible. This observation is also in accordance with the remaining uncertainty regarding inference on the presence of unit roots for this series, as discussed above in Chapter 5. Although the price series for Lamb was ultimately modelled including a trend term, this uncertainty should be borne in mind when interpreting the relative importance of the estimated trend coefficient and the fluctuations around it[8].

In contrast to the model selection undertaken for forecasting purposes in Chapter 5, the price series for Lead is modelled as trend stationary here, for the reasons given above. Again there is considerable volatility surrounding the overall downwards trend, and the illustration provided in Figure 6.4.6 suggests that the impression of a downward trend may in part be attributed to a discrete downwards shift of the series in 1981.

Figure 6.4.6:



[8] Since the alternative of an integrated model may be considered in this case, the Beveridge Nelson decomposition for Lamb is illustrated in Figure VI.iii.i. in Appendix VI.i.

Among the series modelled as stationary for forecasting purposes in Chapter 5, the series of Zinc prices did not include a trend term and is shown as such in Figure 6.4.7 below. In any case, the point estimate for the trend coefficient on the relative price of Zinc was very low and its inclusion would not make a noticeable difference, even over long time horizons. (A plot of the relative price series for zinc and the estimated trend is included in Appendix VI.i.) As was pointed out above, the price series for Coffee will also be shown as a stationary series with the original data and the mean value illustrated in Figure 6.4.8.

Again, both series are subject to substantial volatility overall, with the price series for Zinc showing occasional large deviations with smaller frequent deviations from the mean value.

Figure 6.4.7:

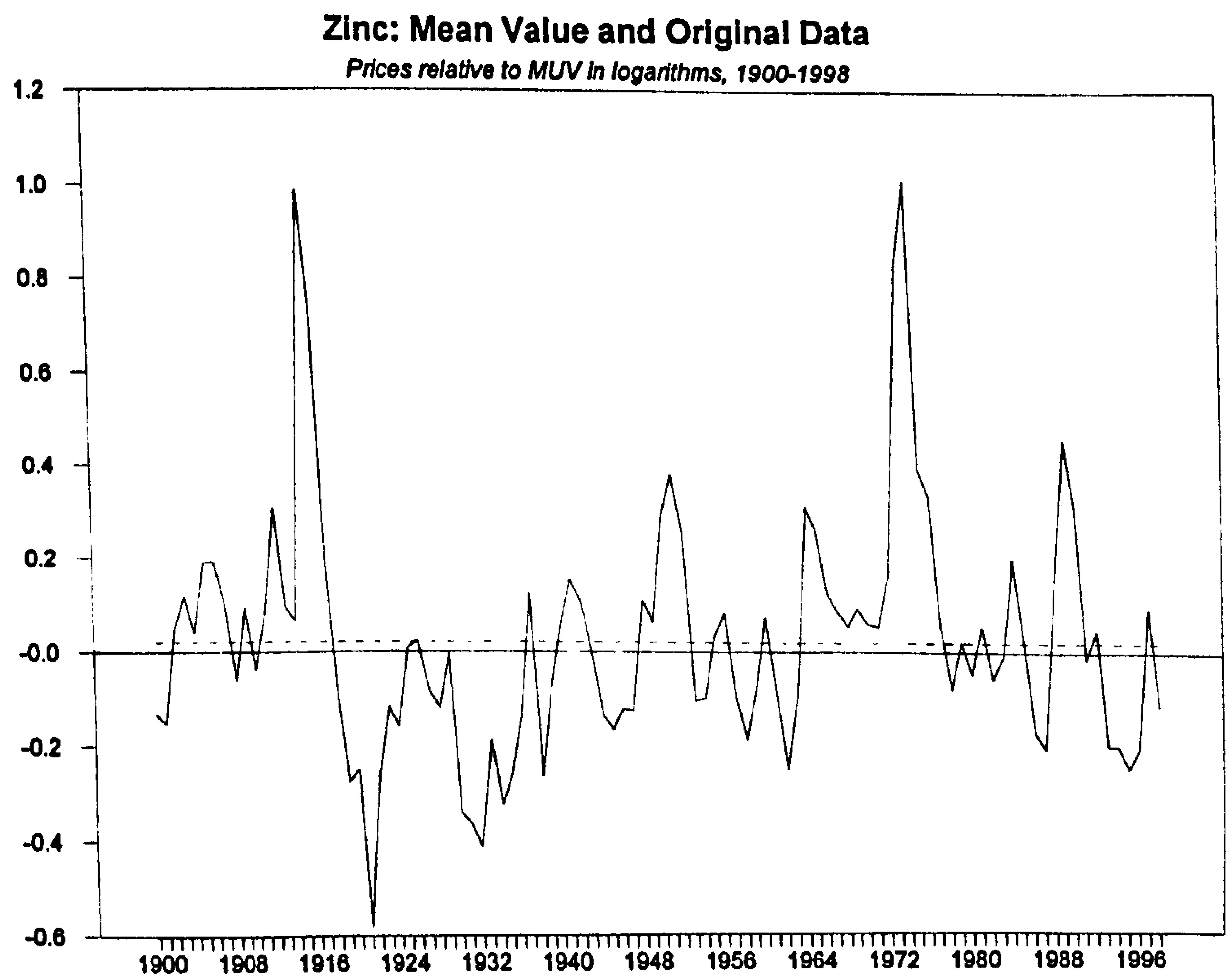
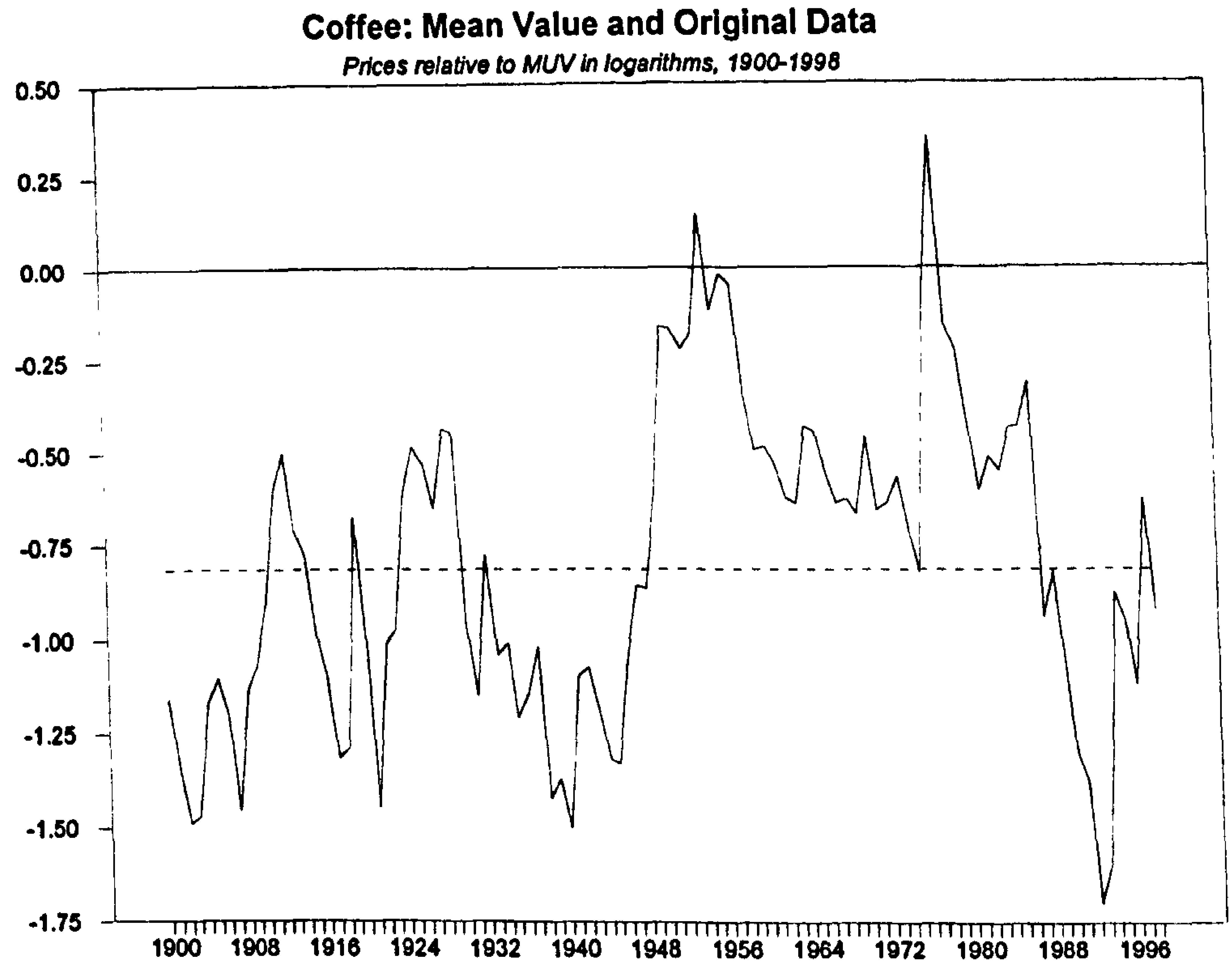


Figure 6.4.8:



6.5. Conclusion: The relative importance trend and cycle components in relative commodity price series

The above exposition reaffirms the previous conclusion that only a limited number of commodity price series are well characterised by a secular trend term. Even in all those cases it can be observed that the volatility surrounding the trend is substantial. The estimated coefficients on the trend term imply an annual change of around 1% for those commodities where a trend term should be considered. (It will be recalled though that the 95% confidence intervals[9] presented in Chapter 3 were rather wide). The associated standard errors of the estimating equation are usually

[9] The confidence intervals in percentage terms where [-1.577, -0.558] for the trend coefficient estimate in the case of Sugar, [1.125, 2.529] for Lamb, [-3.660, -2.015] for Rubber, [0.833, 1.442] for Timber and [-2.279, -1.463] for Aluminium.

much higher taking values around 15-30%[10]. Thus, where the estimated trend coefficients appear to be important in characterising the time path of the data, they have that quality mainly with respect to long run developments. In the short run, the main feature of virtually all the data series under consideration is pronounced volatility.

In the case of the difference stationary models considered, the overall scenario is similar. Inferred shock persistence is usually high, with the permanent data component tending to follow the original data series closely. Furthermore, since the permanent component is itself a random walk or random walk with drift[11], there is also considerable volatility associated with the permanent component. It follows that where a drift term is included and judged to be significant, its estimated value, and its adequacy as a characterisation for the trajectory of the data series over short time horizons, will still be subject to uncertainty.

What an evaluation of trend and cycle components of the data series shows independently of the assumed order of integration of the series is that the most characteristic feature of relative commodity price series over short to medium term horizons is pronounced volatility rather than a clear trend movement in any direction.

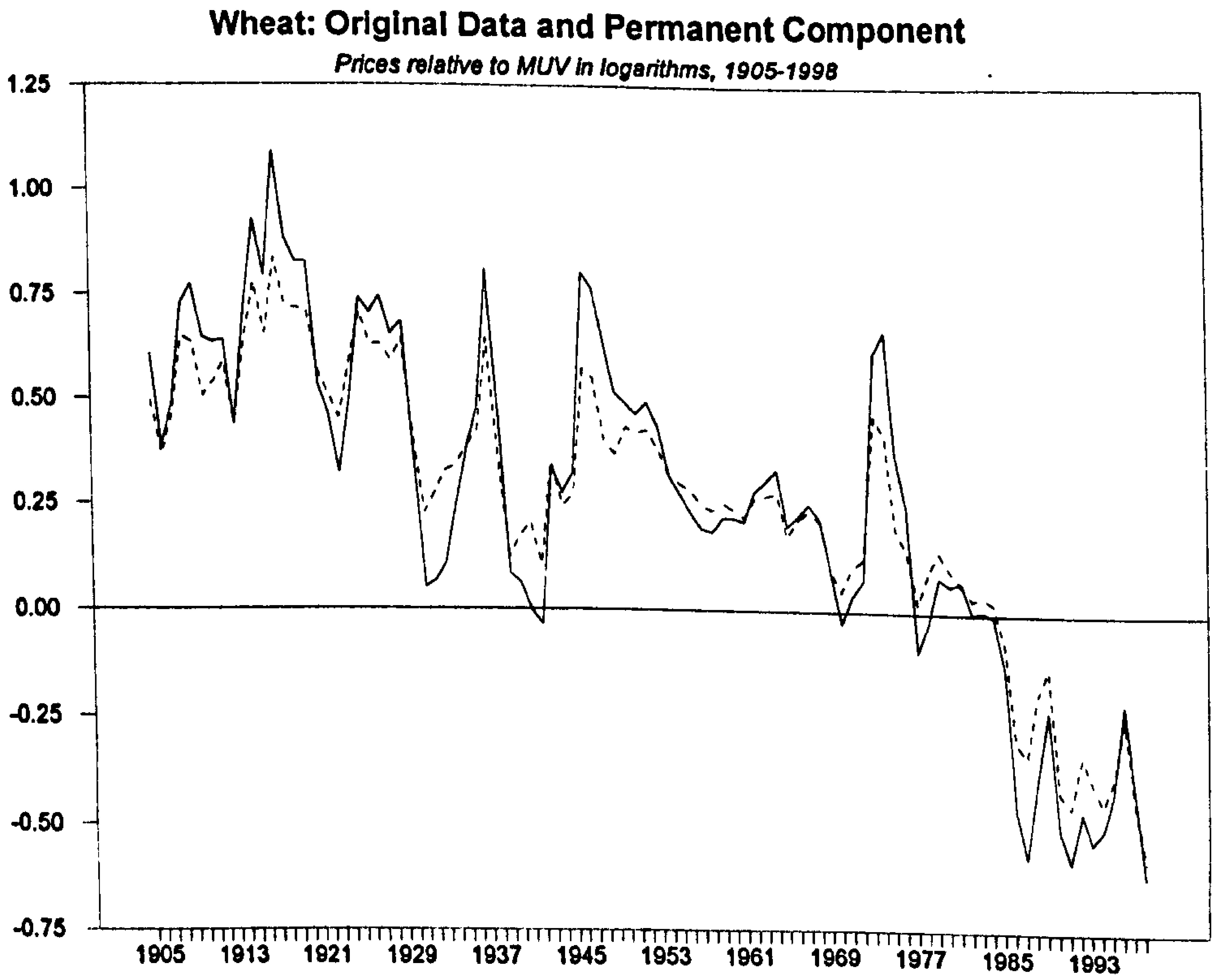
[10]This is evident from the data reported in Appendix III.ii. The trend coefficients and standard errors (*i.e.* $\hat{\sigma}_\varepsilon$) for some commodities are (with standard errors given in parenthesis): -0.011 (0.311) for Sugar, 0.018 (0.194) for Lamb, -0.028 (0.274) for Rubber, 0.011 (0.150) for Timber and -0.019 (0.151) for Aluminium. The estimated trend values and standard errors for Rice and Wheat, where there is some evidence in favour of a significant drift term, are: -0.011 (0.161) for Rice and -0.011 (0.152) for Wheat.

[11] It will be recalled here that the 95% confidence intervals for the drift coefficient estimates (in percentage terms) for Rice and Wheat, as presented in chapter 3 were [-2.136, -0.226] and [-2.637, 0.682] respectively, the drift coefficient for Wheat in the model identified by AIC was [-2.166, 0.021].

Appendix VI.i. Further Graphs of Commodity Price Series and Trend or Permanent Components.

This Appendix presents further illustrations of the data series and permanent or trend components as discussed in Chapter 6. Figure VI.i.i. below shows the original data series and the permanent component of Wheat prices, according to the Beveridge Nelson Definition and when the underlying model is selected using the Akaike Information Criterion (AIC), *i.e.* ARIMA(0,1,4):

Figure VI.i.i:



The Beveridge Nelson permanent component and the original data series for the relative price of Sugar are illustrated in Figure VI.i.ii and of Lamb in Figure VI.i.iii below.

Figure VI.i.ii:

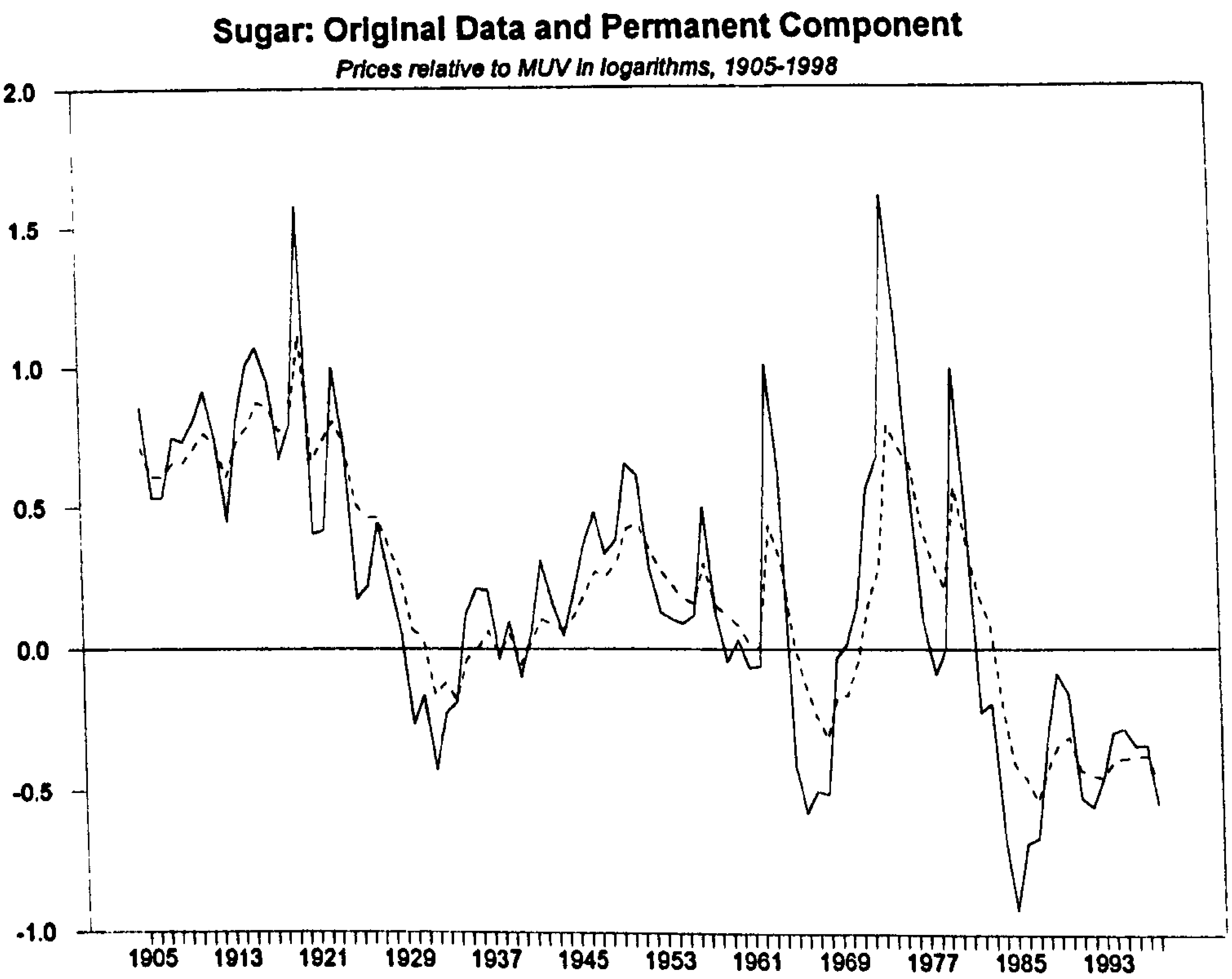


Figure VI.i.iii:

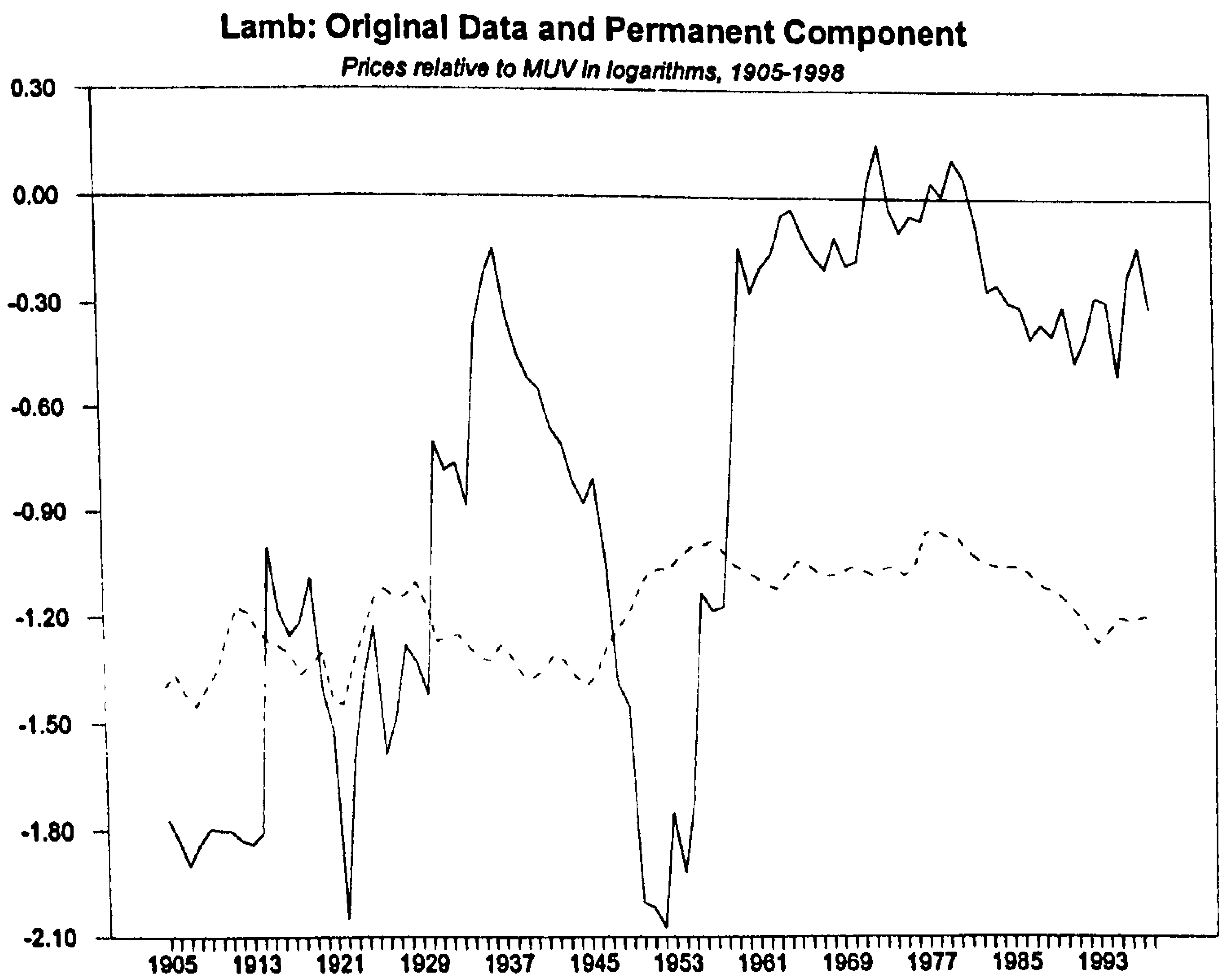
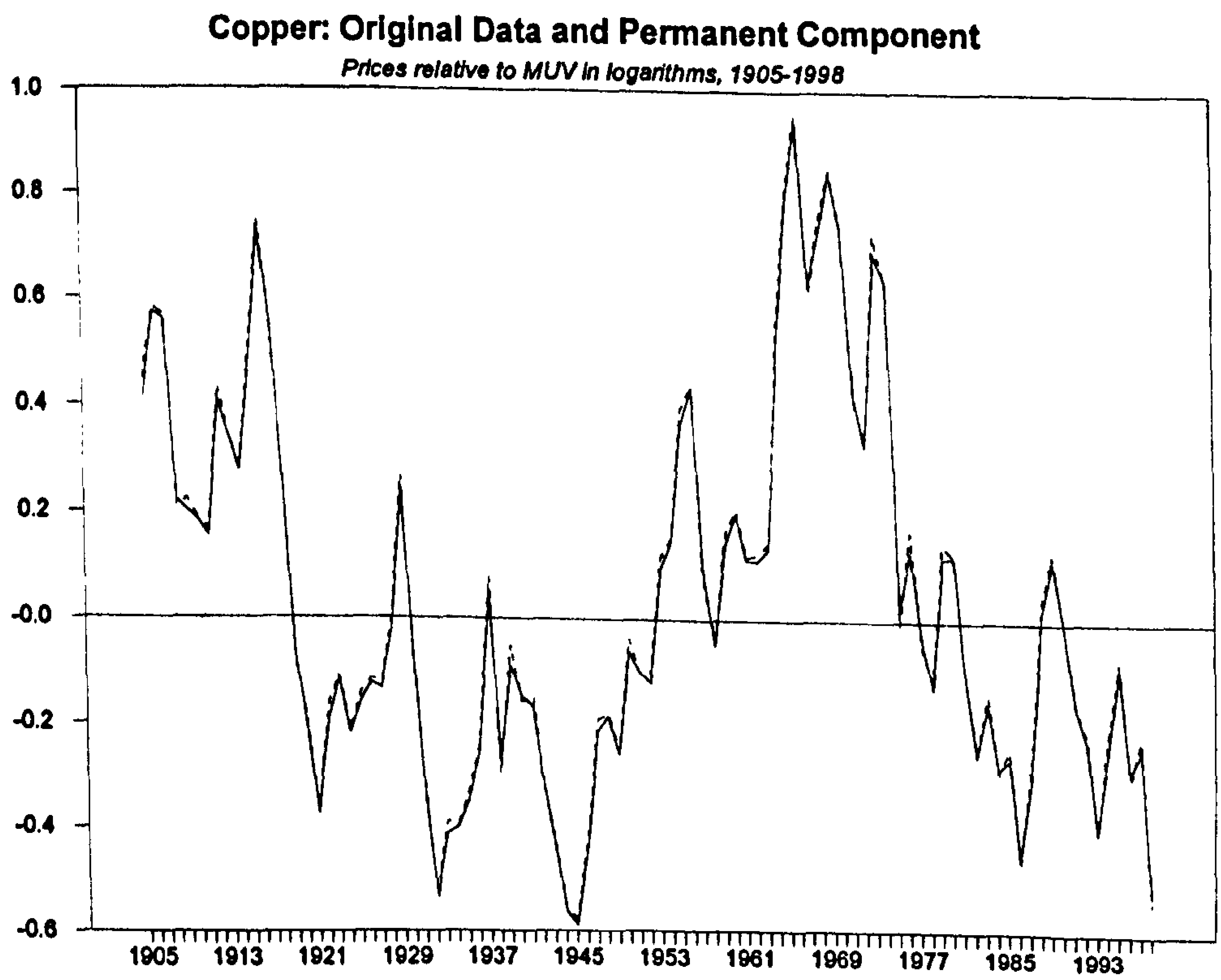


Figure VI.i.iii shows the permanent component and original data series for the price series of Lamb as an alternative to the stationary representation in Chapter 6. The large discrepancy between the permanent component and cycle makes the use of a stationary model seem plausible, as does the estimated value of 0.99999907 for the coefficient on the first order moving average term.

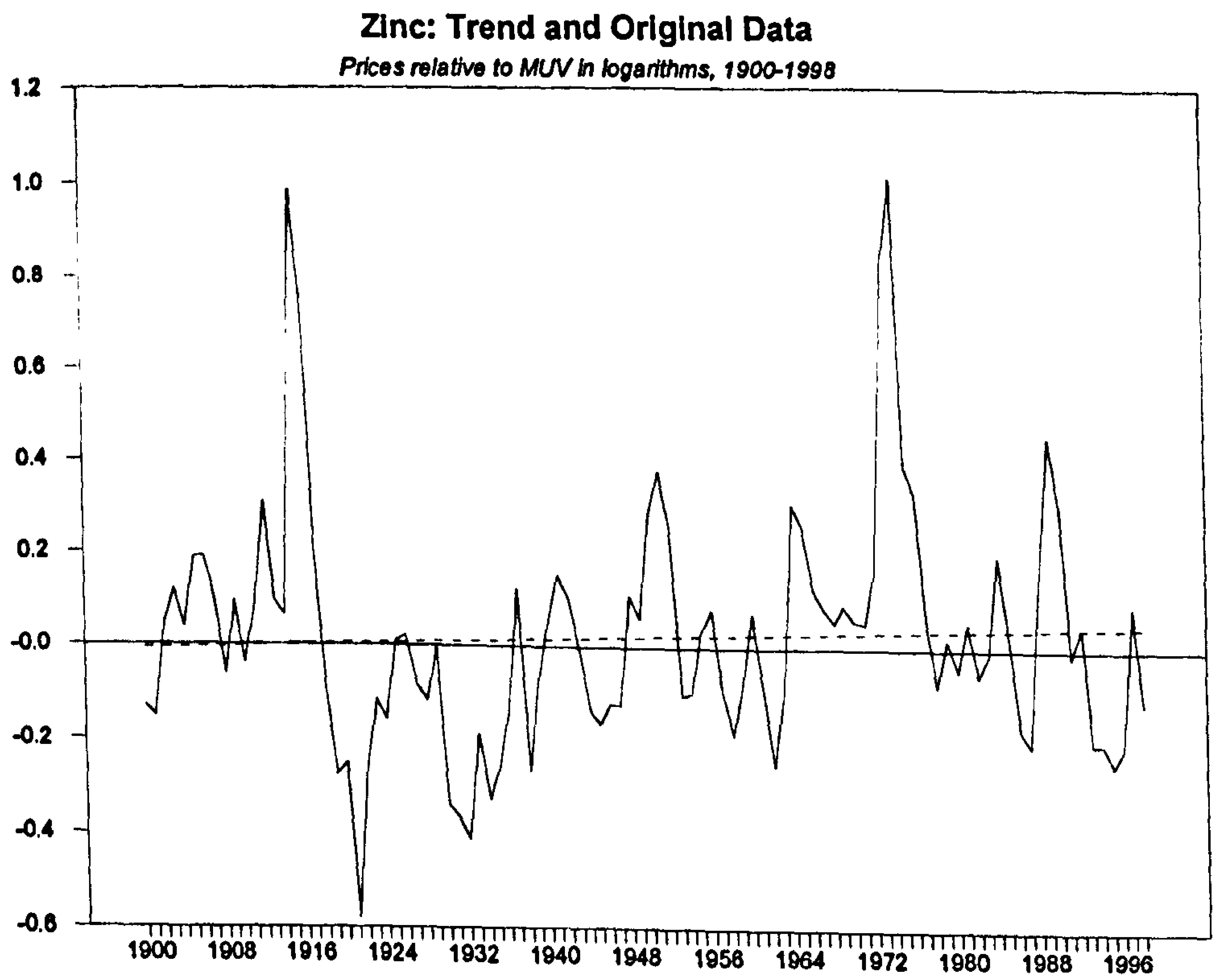
Figure VI.i.iv, for Copper, gives an example for one of several commodity price series which are best characterised as a pure random walk, so that the permanent component and the original data series coincide almost perfectly when an ARIMA(1,1,1) model is imposed on the data.

Figure VI.i.iv:



The small magnitude of the point estimate for the trend on Zinc prices is illustrated in figure VI.i.v below. It is readily appreciated that the trend, if it were significant, would describe only a minimal component of the overall development of the data series, even in the long run.

Figure VI.i.v:



Appendix VI.ii. Persistence Results for Difference Stationary Models

VI.ii.i. Persistence Results

This Appendix details inferred shock persistence results according to [6.2.8]. In each table the proportion of a random disturbance that is taken to be permanent is listed for each commodity together with the number of autoregressive and moving average terms in the corresponding ARIMA model selected by SBC. (An ARIMA(1,1,1) model has been used in those cases where a random walk had been selected as the most appropriate model by SBC).

Table VI.ii.i. shows shock persistence results obtained according to [6.2.8] for those commodity price series where an ARIMA(p,1,q) model with at least $p \neq 0, \text{or}\ q \neq 0$ was selected by SBC[1], and where there is no evidence of overdifferencing. Here and in the following tables, the p autoregressive coefficients and q moving average coefficients are presented in columns three and four respectively.

Table VI.ii.i

Commodity	Persistence	p	q
Cocoa	0.814	2	0
Rice	0.242	1	2
Wheat	0.520	0	2
Maize	0.341	0	2
Sugar	0.431	0	2
Palm Oil	0.765	2	0
Cotton	0.694	2	2
Jute	0.548	0	2
Wool	0.408	0	2
Silver	0.790	2	0

[1] For the ARIMA(0,1,4) model selected by AIC for Wheat the inferred shock persistence would be 0.337.

Table VI.ii.ii. shows the persistence results for those cases where there is evidence of overdifferencing, or a stationary forecast model had been selected in Chapter 5. In those cases where the model selected by SBC was a random walk, an ARIMA(1,1,1) model has been imposed on the data series.

Table VI.ii.ii.

Commodity	Persistence	p	q
Aluminium	0.000	1	2
Coffee	0.000	1	1
Lamb	0.000	1	1
Lead	0.000	1	1
Rubber	0.000	1	1
Sugar	0.431	0	2
Timber	0.000	1	1
Zinc	0.000	1	2

With the exception of Sugar, it can be seen that a zero shock persistence is inferred throughout[2].

Finally, for those series, which are best modelled as a random walk, persistence results obtained when fitting ARIMA(1,1,1) models are shown in table VI.ii.iii.

Table VI.ii.iii

Commodity	Persistence	p	q
Bananas	1.075	1	1
Beef	1.049	1	1
Copper	1.073	1	1
Tea	1.009	1	1
Tin	1.060	1	1
Tobacco	1.070	1	1

[2] Since the persistence results listed in the tables are not given in percentage terms, the inferred persistence for Lead rounds to zero

Here the inferred shock persistence is in excess of 100% in all cases. This is most appropriately explained by reference to the random walk models originally selected as appropriate representations of the data series.

VI.ii.ii. A comparison with the persistence results of Cuddington (1992)

Following the comparison in Chapter 3 of trend results with those obtained by Cuddington (1992) a further comparison is made here between the persistence results obtained here and those of Cuddington (*op. cit.*)[3]. By its very nature, this comparison, is restricted to those commodities that have been modelled as difference stationary in Cuddington (*op. cit.*). Table VI.ii.iv lists the persistence results of Cuddington (*op. cit.*) and those obtained here according to [6.2.8] for the relevant commodities.

Table VI.ii.iv Comparison with Persistence Results from Cuddington (1992)

Commodity	Cuddington (1992)	[6.2.8]*
Aluminium	0.883	0.000
Bananas	0.444	1.000
Beef	1.000	1.000
Cocoa	0.641	0.814
Copper	1.000	1.000
Cotton	0.561	0.694
Jute	0.400	0.548
Rubber	1.000	0.000
Silver	0.641	0.790
Tea	0.720	1.000
Tobacco	0.781	1.000
Wool	0.343	0.408

* Persistence computed according to [6.2.8] in Chapter 6.

In those cases, where a random walk has been found to be the most appropriate representation of the data series, its persistence in Table VI.ii.iv is shown as 1 (or

[3] León and Soto (1997) also look into the issue of shock persistence but use a variance ratio statistic to assess mean reversion and shock persistence over long time horizons. Since the length of the time horizon covered extends over half the sample size (or about 49 years in the present case) the period considered for the assessment of the transitory component by León and Soto (*op.cit.*) is considered too long, given the context of short to medium term forecasts in the present study.

100%) rather than the value corresponding to the overfitted ARIMA(1,1,1) model which had been listed in Table VI.ii.iii.

Except for those cases where a random walk has been selected as the appropriate model in both studies (*i.e.* Beef and Copper) the inferred persistence results are not exactly equal. Given that the data sets used are of different length and that the selected model parameterisations are not always identical, this is to be expected. What is remarkable are the different persistence results obtained for Aluminium and Rubber.

These results are likely to be related, not only to the different length of the sample, but also to differing inferences on unit roots and different model specifications.

For Aluminium the selected difference stationary model specifications differ between Cuddington and the present Study. (Cuddington uses an ARIMA(0,1,5) model in which he dropped the third and fourth moving average term). The data series also has been much closer to the estimated trend line during the later part of the sample period, which had not been included in Cuddington's case. Thus, differences in the selected models and in the sample data may well be important factors of influence here.

In the case of Rubber, the difference stationary model selected by SBC is a random walk, as in Cuddington (*op. cit.*). Had persistence been inferred for the difference stationary model for Rubber thus specified, the results would have been identical to those of Cuddington. Since, in light of the evidence on possible overdifferencing presented in Chapter 4, persistence results have been computed for the case of an

ARIMA(1,1,1) model, the inferred low shock persistence that should be expected for a trend stationary process has been obtained.

This case then confirms the disadvantage of relying on discrete *a priori* decisions regarding the presence of unit roots in moderately sized samples. (Cuddington did consider the possibility of non invertible moving average processes as an indicator of overdifferencing, and thus did not rely on ADF test results alone.) It is particularly in cases like the one of Rubber, that simulation results on the finite sample impact of serially correlated residuals on trend coefficient hypothesis tests can provide helpful complementary evidence.

Chapter 7

Overall Conclusions

Chapter 7: Overall Conclusions

This study has investigated the presence of trends in relative primary commodity price series for individual commodities, using trend stationary and difference stationary univariate models, and has projected these price series over a period of ten years. The data used were those underlying the Grilli and Yang index, extending the time period covered up to 1998 for most commodities. Conclusions on the presence of trend or drift coefficients were found to depend crucially on the inferred order of integration. The explicit consideration of structural breaks, however, did little to clarify the results.

It has been shown that a drop in the significance level of trend coefficient estimates can occur in univariate data series when a difference stationary representation is used for a trend stationary process. This problem was linked to the occurrence of underparameterised difference stationary model alternatives and it was shown that the problem can be alleviated by either setting a minimum number of autoregressive and moving average lags for the ARMA(p,q) residual parameterisation or by preferring the Akaike Information Criterion (AIC) over the Schwarz Bayesian Criterion (SBC) if model selection is undertaken by information criterion. In the particular case where a random walk, allowing for the presence of a drift, is considered as an alternative to an ARIMA(1,0,0) model with trend, underspecification should be detected by fitting an ARIMA(1,1,1) model as a difference stationary alternative.

The contrary result of spurious rejections of the null hypothesis of a zero trend coefficient when a stationary model is fitted to an integrated data series is a well documented phenomenon (*cf.* Newbold and Granger (1974)). It was here confirmed that the possibility of spurious rejections is still a problem in finite samples where serial correlation is taken into account. Simulation experiments were undertaken to quantify the finite sample impact of serial correlation in the present cases and empirical critical values were obtained for some of the models estimated. These were then used to assess the significance of trend estimates obtained while considering the impact of serial correlation. In the case of one commodity, which could be classified as stationary (*viz* Lead) it was then shown that evidence in favour of a significant trend coefficient is substantially weakened if the impact of serial correlation is taken into account.

In addition, a test for the significance of trend coefficients developed by Vogelsang (1998) was applied to the data series in question and was further tested in Monte Carlo simulations representing the alternative model parameterisations considered here. The test indicated the presence of a trend term in few of the series where a trend has been suggested by the other methods employed. Further simulations, examining the performance of the test for the different trend stationary and difference stationary model alternatives considered in this study did not confirm the Vogelsang test as a reliable indicator of trend components in stationary or difference stationary series. The Vogelsang test seems to overcome the problem of spurious rejections of the null hypothesis of a zero trend coefficient, but the test has been shown to have low power, in particular where the underlying series is integrated.

In conclusion, strong support for the presence of a trend or drift coefficient was found for only seven commodity price series: Aluminium, Hides, Rice, Wheat, Rubber, Sugar and Timber. Weaker evidence has been obtained for the presence of a trend term in the series for Lamb. Thus, the hypothesis of a statistically significant trend or drift term can be supported for up to eight of the 24 relative commodity price series considered. Six of the eight trend and drift coefficient estimates in question are negative. In particular, a secular decline is most likely for Aluminium, Hides, Rubber and Sugar, followed by Rice and Wheat. The negative trend estimates imply a decline of around 1.0 to 1.2 percent per year, except for Aluminium and Rubber where the estimated magnitude is of 1.9 percent and 2.8 to 3 percent respectively. Positive trend coefficients have been inferred for Lamb and Timber. The estimated rise in the relative price of Lamb is of 1.8 to 1.5 percent per year while the corresponding estimate for Timber takes a value of 1.1 to 0.8 percent per year.

Regarding forecast model selection, it was shown by Diebold and Kilian (2000) that improved forecasts can be obtained when determining the order of integration of the forecast model through ADF tests. It was here investigated whether these results continue to hold if inference on the presence of a trend term is itself conditional upon the inferred order of integration of the series. It was shown that, in the case of a trend stationary first order autoregressive data generating process, consistently imposed differenced forecast models can outperform models selected by pre-testing if the trend coefficient value is low and if the value of the AR(1) coefficient is sufficiently large. (This result implies that in some cases superior forecast results were obtained from consistently applied difference stationary forecast models, at lower AR(1) coefficient values than would have been the case if the presence of a trend coefficient had been correctly inferred in either case.) Additional simulation results confirmed that a lower cost of misspecification is generally associated with difference stationary forecast models. It was also shown that trend and drift coefficients only tend to improve forecast results from a certain minimum magnitude, and that the required magnitude for a drift term is higher than that for a trend term in a trend stationary forecast model. Trend terms were incorporated into the forecast models for Aluminium, Rubber, Timber, Sugar and Lamb. Difference stationary forecast models with a drift term have been selected for Rice and Wheat, while the series for Zinc has been the only stationary price series forecast without a trend term. Short term forecasts are often dominated by the ARMA(p,q) parameterisation of the residual process. Wheat prices for example were predicted to rise temporarily before falling consistently in line with the negative drift estimate.

In addition to the forecast results obtained, shock persistence results were calculated on the basis of gain functions and a decomposition into transitory and permanent components was undertaken using the Beveridge Nelson Decomposition. The decomposition results obtained in this way are consistent with the shock persistence results obtained from the gain function. The inferred shock persistence is often high, although low shock persistence results were obtained for price series where a stationary or trend stationary representation is likely to be appropriate.

The hypothesis of a generally negative trend or of generally homogeneous time series characteristics for relative primary commodity prices could not be confirmed in this study. It was concluded though that inference on the presence of trend components in the data series is more strongly influenced by inference on the order of integration of the data series than by attempts to account for structural breaks. While conclusions on the presence of the trend could at times be reached independently of stationarity assumptions and while some of the problems associated with modelling integrated time series with drift could be corrected, the inferred order of integration was found to remain important in many cases. Negative significant trend estimates were found to be more frequent than significant positive trend estimates, although the hypothesis that relative primary commodity prices are generally characterised by negative secular trends or even secular trends *per se* can not be upheld. Of the eight trend or drift coefficient estimates which are likely to be statistically significant six are negative. (Considering point estimates of trend coefficients without regards to significance, 16 out of 24 coefficients[1] take negative signs, 8 do not.)

Given the differing conclusions on the presence of trend coefficients in the series, it should come as no surprise, that the forecasts obtained also tend to differ. Extensive price volatility appears to be one of the common features of the

[1] It will be recalled, that the series for Hides was considered in Chapter 4 only, since no updated figures were available after 1995.

commodity price series covered. Inference on shock persistence however differs substantially. While for some series shocks were classified as mean reverting there is substantial shock persistence for others.

Given the different time series characteristics of the various relative commodity price series in a number of respects, one conclusion suggested by this study is that one should be reluctant to generalise with respect to the time series properties of primary commodity prices. This casts doubt on the usefulness of aggregate commodity price indices frequently used in empirical work on this topic. Furthermore, it appears that generalisations about primary commodity price behaviour can be misleading even for groups of related commodities, such as metals, cereals or tropical beverages. (The opposite conclusions reached on the presence of a drift term in the price series for Wheat and Maize, for example, highlight this point, as does the marked difference in the long run behaviour of Aluminium and Copper prices among metals.)

The implied implausibility of generalised assumptions regarding developing countries' barter terms of trade based on the time series properties of relative primary commodity prices is likely to complicate the analysis of the long term behaviour of these terms of trade. There are even fewer long run data series available on developing countries' terms of trade than on commodity prices (see for example León and Soto (1995)). A good understanding of the time series behaviour of individual price series however may still contribute towards understanding the long run properties of terms of trade series if the main participating export commodities are known. How far this conclusion carries over to predictions of terms of trade series, will partly depend on how far the evolution of the sectoral composition of developing countries' trade profiles can be anticipated.

The assumption that relative commodity prices are best described as generally following a negative trend is not supported by the evidence in this study. In how far a careful study of individual price series can complement the available data series for individual developing countries' terms of trade is a subject for further investigation.

Bibliography

Bibliography

Agiakloglou, Christos, Newbold, Paul (1996) *The balance between size and power in Dickey-Fuller tests with data dependent rules for the choice of truncation lag*, Economics Letters, Vol. 52, No.3, pp. 229-234

Agiakloglou, Christos and Paul Newbold (1992) *Empirical Evidence on Dickey-Fuller Type Tests*, Journal of Time Series Analysis, Vol. 13, No.6, pp.471-483

Ahrens, Ashley and Vija Sharma (1997) *Trends in Natural Resource commodity Prices: Deterministic or Stochastic?*, Journal of Environmental Economics and Management, Vol.33, No.1, pp.59-74

Anderson, Kym (1987) *On why agriculture declines with economic growth*, Agricultural Economics, Vol. 4, No.1, pp.195-207

Antle, John (1988) *World Agricultural Development and the Future of U.S. Agriculture*, American Enterprise Institute, Washington D.C.

Ardeni, Pier Giorgio and Brian Wright (1992) *The Prebisch-Singer Hypothesis: A Reappraisal Independent of Stationarity Hypotheses*, The Economic Journal, Vol. 102, No.413, pp. 803-812

Beveridge, Stephen, and Charles R. Nelson (1981) *A new Approach to Decompostion of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'.*, Journal of Monetary Economics, Vol. 7, No. 2, pp. 151-174

Bleaney, Michael and David Greenaway (1993) *Long-Run Trends in the Relative Price of Primary Commodities and in the Terms of Trade of Developing Countries*, Oxford Economic Papers, Vol. 45, No.3, pp. 349-363

Bloch, Harry and David Sapsford (2000) *Whither the terms of trade? An elaboration of the Prebisch-Singer hypothesis*, Cambridge Journal of Economics, Vol. 24, No. 4, pp. 461-481

Bloch, Harry, Sapsford, David (1997) *Some Estimates of Prebisch and Singer Effects on the Terms of Trade between Primary Producers and Manufacturers*, World Development, Vol. 25, No. 11, pp. 1873-1884

Bloch, Harry, Sapsford, David, (1991-92) *Postwar movements in prices of primary products and manufactured goods*, Journal of Post Keyenesian Economics, Vol. 14, No. 2, pp. 249-266

Borzenstein, Eduardo and Carmen Reinhart (1994) *The Macroeconomic Determinants of Commodity Prices*, IMF Staff Papers, Vol. 41, No.2, pp. 236-261

Borzenstein, Eduardo, Mohsin S. Khan, Carmen M. Reinhart and Peter Wickham (1994) *The Behavior of Non-Oil Commdodity Prices*, IMF Occasional Paper, No. 112

Box, George, Jenkins, Gwilym, (1976) *Time Series Analysis -forecasting and control*, Holden Day, San Francisco

Cashin, Paul and C. John McDermott (2001) *The Long-Run Behavior of Commodity Prices: Small Trends and Big Variability*, Working Paper of the International Monetary Fund, WP/01/68

Cashin, Paul, and Catherine Patillo (2000) *Terms of Trade Shocks in Africa: Are They Short-Lived or Long-Lived?*, Working Paper of the International Monetary Fund, WP/00/72

Cashin, Paul, McDermott, John, Alasdair Scott (1999a) *The Myth of Comoving Commodity Prices*, Working Paper of the International Monetary Fund, WP/99/169

Cashin, Paul, C. John McDermott and Alasdair Scott (1999b) *Booms and Slumps in Primary Commodity Prices*, Working Paper of the International Monetary Fund, WP/99/115

Clements, Michael and DF Hendry (2001) *Forecasting with difference stationary and trend-stationary models*, Econometrics Journal, Vol. 4, No.1 pp. S1-S19

Cuddington, John, (1992) *Long-run trends in 26 primary commodity prices*, Journal of Development Economics, Vol 39, No.2, pp. 207-227

Cuddington, John and Carlos M. Urzúa (1989) *Trends and cycles in the net barter terms of trade: a new approach*, The Economic Journal, Vol. 99, No.396, pp. 426-442

Cuddington, John and Alan Winters (1987) *The Beveridge-Nelson Decomposition of Economic Time Series*, Journal of Monetary Economics, Vol.19, No.1, pp.125-127

Diebold, Francis and Lutz Kilian (2000) *Unit-Root Tests are Useful for Selecting Forecasting Models*, Journal of Business and Economic Statistics, Vol.18, No.3, pp.265-273

Enders, Walter, (1995) *Applied Econometric Time Series*, Wiley Series in Probability and Mathematical Statistics, John Wiley&Sons Inc, New York

Gandolfo, Giancarlo (1994) *International Economics I*, Springer Verlag, Berlin

Granger, Clive and Paul Newbold (1986) *Forecasting Economic Time Series*, second edition, Academic Press

Grilli, Enzo, Yang, Maw Cheng, (1988) *Primary Commodity Prices, Manufactured Goods Prices, and the Terms of Trade of Developing Countries: What the Long Run Shows*, The World Bank Economic Review, Vol. 2, No.1, pp.1-47

Harvey, Andrew (1993) *Time Series Models*, Harvester Wheatsheaf

Helg, Rodolfo (1991) *A note on the stationarity of the primary commodities relative price index*, Economics Letters, Vol.36, No.1, pp.55-60

Ingersent, Ken A., Anthony J. Rayner (1999) *Agricultural Policy in Western Europe and the United States*, Edward Elgar Publishing

Johnston, Jack and John Dinardo (1997) *Econometric Methods*, McGraw Hill, 4th edition

Kaplinsky, Raphael (1999) *"If you want to get somewhere else, you must run at least twice as fast as that!": The Roots of the East Asian Crisis*, Competition & Change, Vol.4, No.1, pp.1-30

Kindleberger, Charles (1958) *The Terms of Trade and Economic Development*, Review of Economic and Statistics, Vol.40, No.1, pp 72-85

Kwiatkowski, Denis, P.C.B. Phillips, P. Schmidt and Y. Shin (1992) *Testing the null hypothesis of stationarity against the alternative of a unit root*, Journal of Econometrics, Vol.54, No.1-3, pp.159-178

León Javier and Raimundo Soto (1997) *Structural Breaks and Long-Run Trends in Commodity Prices*, Journal of International Development, Vol.9, No.3, pp.347-366

León Javier and Raimundo Soto (1995) *Términos de intercambio en la América Latina, Una cuantificación de la hipótesis de Prebisch y Singer*, El Trimestre Económico, Vol.62, No.2, pp.171-199

Leybourne, Steve and B.P.M. McCabe (1999) *Modified Stationarity Tests with Data-Dependent Model-Selection Rules*, Journal of Business & Economic Statistics, Vol.17, No.2, pp.264-270

Leybourne, Steve and B.P.M. McCabe (1994) *A Consistent Test for a Unit Root*, Journal of Business & Economic Statistics, Vol. 12, No.2, pp.157-168

Lutz, Matthias (1999a.) *A General Test of the Prebisch-Singer Hypothesis*, Review of Development Economics, Vol.3, No.1, pp.44-57

Lutz, Matthias (1999b.) *Commodity Terms of Trade and Individual Countries' Net Barter Terms of Trade: Is There an Empirical Relationship?*, Journal of International Development, Vol.11, No.6, pp.859-870

Lutz Matthias and Hans W. Singer (1994) *Trend and Volatility in the Terms of Trade: Consequences for Growth*, in: Sapsford, D., Morgan, W., (Eds.), Economics of Primary Commodity Prices, Edward Elgar, Aldershot

Maizels, Alfred, (1992) *Commodities in Crisis*, Clarendon Press, Oxford

Myrdal, Gunnar, (1989) *Trade as a Mechanism of International Inequality*, in: Leading Issues in International Development, Oxford University Press, Oxford

Newbold, Paul, Anthony Rayner and Neil Kellard, (2000) *Long-run Drift, Co-Movement and Persistence in Real Wheat and Maize Prices*, Journal of Agricultural Economics, Vol.51, No.1, pp.106-121

Newbold, Paul, Vougas, Dimitrios, (1996) *Drift in the relative price of primary commodites: a case where we care about unit roots*, Applied Economics, Vol. 28, No.6, pp. 653-661

Newbold, Paul (1995) *Statistics for Business and Economics*, Prentice Hall

Newbold Paul (1991) *Structural Decomposition of Time Series with Implications in Economics, Accounting, and Finance Research*, Review of Quantitative Finance and Accounting, Vol. 1, No.3, pp. 259-279

Newbold, Paul (1990) *Precise and efficient computation of the Beveridge-Nelson decomposition of economic time series*, Journal of Monetary Economics, Vol. 26, No.3, pp. 453-457

Newbold, Paul and Clive Granger (1974) *Spurious Regressions in Econometrics*, Journal of Econometrics, No.2, Vol.2, pp.111-120

Perron, P., (1989) The great crash, the oil price shock and the unit root hypothesis, Econometrica, Vol.57, No.6, pp. 1361-1401

Powell, Andrew (1991) *Commodity and Developing Country Terms of Trade: What Does the Long Run Show?*, The Economic Journal, Vol.101, No.409, pp.1485-1496

Prebisch, Raúl (1959) *Commercial Policy in the Underdeveloped Countries*, American Economic Review, Vol. 49, No.2, pp. 251-273

Reinhart, Carmen M. and Peter Wickham (1994) *Commodity Prices: Cyclical Weakness or Secular Decline?*, IMF Staff Papers, Vol. 41, No. 2, June 1994

Ricardo, David (1951) *On the Principles of Political Economy and Taxation*, Ch.VII, The Works and Correspondence of David Ricardo, Vol. 1, Cambridge University Press, Cambridge

Sapsford, David and Hans Singer (1998) *The IMF, the World Bank and Commodity Prices: A Case of Shifting Sands?*, World Development, Vol.26, No.9, pp.1653-1660

Sapsford, David, P., Sarkar and Hans Singer (1992) *The Prebisch-Singer Terms of Trade Controversy Revisited*, Journal of International Development, Vol.4, No.3, pp.315-332

Sapsford, David (1985) *The Statistical Debate on the Net Barter Terms of Trade Between Primary Commodities and Manufactures: A Comment and Some Additional Evidence*, The Economic Journal, Vol. 95, No.379, pp. 781-788

Sarkar, Prabirjit (1997) *Growth and the Terms of Trade: A North-South Macroeconomic Framework*, Journal of Macroeconomics, Vol. 19, No. 1, pp. 117-133

Singer, Hans W. (1975) *The Distribution of Gains Revisited*, In: The Strategy of International Development, Singer (ed) pp 58-66, Macmillan Press London

Singer, Hans W. (1958) *The Terms of Trade and Economic Development: Comment*, The Review of Economics and Statistics, Vol.40, No.1, pp.85-90

Singer, Hans W. (1950) *U.S. Foreign Investment in Underdeveloped Areas -the Distribution of Gains between Investing and Borrowing Countries*, American Economic Review -Papers and Proceedings, Vol.40, No.2, pp. 473-485

Spraos, John, (1980) *The Statistical Debate on the Net Barter Terms of Trade Between Primary Commodities and Manufactures*, The Economic Journal, Vol. 90, No.9, pp. 107-128

Stock, James H. (1996) *VAR, Error Correction and Pretest Forecasts at Long Horizons*, Oxford Bulletin of Economics and Statistics, Vol.58, No.4, pp.685-701

Sun, Hongguang, Pantula, Sastry, (1999) *Testing for trends in correlated data*, Statistics & Probability Letters, Vol. 41, No.1, pp. 87-95

Vial Joaquín (1992) *Copper consumption in the USA*, Resources Policy, Vol. 18, No.2, pp.107-121, June 1992

Von Hagen, Juergen (1989) *Relative Commodity Prices and Cointegration*, Journal of Business and Economic Statistics, Vol.7, No.4, pp.497-503

Vogelsang, Timothy, (1998) *Trend Function Hypothesis Testing in the Presence of Serial Correlation*, Econometrica, Vol. 66, No.1, pp. 123-148

Wood, Adrian (1994) *North-South Trade Employment and Inequality*, Clarendon Press, Oxford

Worldbank (2000) *Commodities in the 20th Century*, Global Commodity Markets, January 2000
http://www.worldbank.org/prospects/gcmonline/01specialfeature.pdf

Data Sources:

Pinksheets:

Worldbank (2001) Global Economic Prospects and the Developing Countries - Commodity Price Data / Pinksheet, March 2001

Worldbank (1999) Global Economic Prospects and the Developing Countries - Commodity Price Data / Pinksheet, July 1999

Worldbank (1998) Global Economic Prospects and the Developing Countries - Commodity Price Data / Pinksheet, January 1998

World Development Indicators:

Worldbank (2000) World Development Indicators, Table 6.4 Primary Commodity Prices

Worldbank (1998) World Development Indicators, Table 6.5 Primary Commodity Prices

Not all data could be obtained from the above data sources and some data for Wool, Lamb and Timber were supplied by Betty Dow from World Bank sources.